

PR 27-JUL-2000; 2000US-0221670P.
 PR 07-AUG-2000; 2000US-0223170P.
 PR 07-AUG-2000; 2000US-0223172P.
 PR 07-AUG-2000; 2000US-0223460P.
 PR 26-OCT-2000; 2000US-0244037P.
 PR 26-OCT-2000; 2000US-0244111P.
 PR 26-JUN-2001; 2001US-0301217P.
 (DELT-) DELTAGEN INC.
 XX
 PI Allen KD, Leviten MW;
 XX MPI; 2002-154953/20.
 DR P-PSDB; AAE23083.
 PT Novel non-human transgenic animal, preferably transgenic mice comprising
 PT disruption in target gene, e.g., trypsinase gene, useful for identifying an
 agent that modulates expression or function of target gene.
 PS Example 2; Fig 2A; 74pp; English.
 XX
 CC The present invention relates to non-human transgenic animals preferably
 CC transgenic mice comprising disruption in target gene such as trypsinase
 CC gene. The invention also relates to compositions and methods relating to
 CC the characterization of gene functions. The transgenic animals are useful
 CC for identifying an agent that modulates the expression or function of a
 CC target. They are useful for identifying an agent that modulates a
 CC phenotype associated with a disruption in trypsinase genes or limulus
 CC clotting factor protease-like genes by administering an agent to the
 CC transgenic animal and determining whether the agent modulates the
 CC phenotype where the agent has effect on decreased body weight, decreased
 CC thymus weight, decreased thymus to body weight ratio, increased pre-pulse
 CC inhibition, significant decrease in their response latency to the hot
 CC plate test or a decreased response threshold to metrazol. Agents that
 CC modulate the expression, function or activity of the target gene are
 CC useful for treating a disorder associated with a mutation in trypsinase
 CC gene or in limulus clotting factor protease-like gene. The transgenic
 CC animals are useful for testing the efficacy of proposed genetic and
 CC pharmacological therapies for human genetic diseases. They are useful as
 CC models for diseases, disorders or conditions associated with phenotypes
 CC relating to a disruption in a target and to identify drugs,
 CC pharmaceuticals, therapies and interventions which may be effective in
 CC treating a disease or other phenotypic characteristics of the animal. The
 CC present sequence is epithin gene. This sequence is used in the
 CC exemplification of the invention
 XX
 SQ Sequence 3106 BP; 692 A; 862 C; 897 G; 655 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3106; DB 6; Length 3106;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 301 ATTATCGAATGTGGGGTTCAAAAAGTCTTCATGCGCATCGAGATCAAAATGAGA 360
 Qy 361 TCTTTCTGATGCGTATGGAATCTCCACCTCCACAGATTATCAGCTGGCCAGCAG 420
 Db 361 TCTTTCTGATGCGTATGGAATCTCCACCTCCACAGATTATCAGCTGGCCAGCAG 420
 Qy 421 TGAAGAGCGCTGAAGCTGCTGTACATGAAGTCCCTGCTGGGTCCCTACCAAGA 480
 Db 421 TGAAGAGCGCTGAAGCTGCTGTACATGAAGTCCCTGCTGGGTCCCTACCAAGA 480
 Qy 481 AGTGGCTGTAACTGCTTCACTGAGAGGCAAGTCACTGCTTCTACTGCTGAGTTCA 540
 Db 481 AGTGGCTGTAACTGCTTCACTGAGAGGCAAGTCACTGCTTCTACTGCTGAGTTCA 540
 Qy 541 GCATCCCCCAACCTGGCAGAGAGTTGATGCGGCATAGGCTGTGAGACGAGTTTAA 600
 Db 541 GCATCCCCCAACCTGGCAGAGAGTTGATGCGGCATAGGCTGTGAGACGAGTTTAA 600
 Qy 601 CATTTGCCACCCCGAGGACCGGGCACTGAATCTTCTGCTTACATCTGTGGTGGCTTCC 660
 Db 601 CATTTGCCACCCCGAGGACCGGGCACTGAATCTTCTGCTTACATCTGTGGTGGCTTCC 660
 Qy 661 CCAATTGACCCCAAGATGCTTCAAGAGATCTGAGAGCAACAGTGAATTTGCCCTGATG 720
 Db 661 CCAATTGACCCCAAGATGCTTCAAGAGATCTGAGAGCAACAGTGAATTTGCCCTGATG 720
 Qy 721 CCATGGTGTGAGAGATGACACGCTTCACTACCCCTGCTTCCCAAGATCCCTACCCGG 780
 Db 721 CCATGGTGTGAGAGATGACACGCTTCACTACCCCTGCTTCCCAAGATCCCTACCCGG 780
 Qy 781 CGCATGCTCGCTGCGCAATGGGTCTGTGCGGGGAGCGCGACCTGTGCTGAGCTTCACT 840
 Db 781 CGCATGCTCGCTGCGCAATGGGTCTGTGCGGGGAGCGCGACCTGTGCTGAGCTTCACT 840
 Qy 841 TCCGAAGCTTGAATGTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATG 900
 Db 841 TCCGAAGCTTGAATGTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATG 900
 Qy 901 ATAGCTGAGCGCCCATGAAACCCCAAGCTGTGTGCGGCTGTGTGAGCACTTCTCACCT 960
 Db 901 ATAGCTGAGCGCCCATGAAACCCCAAGCTGTGTGCGGCTGTGTGAGCACTTCTCACCT 960
 Qy 961 CCTAACCACTGACTTCTCTCTCTCTCCAGAAAGCTTCTTGTGTGATGATGATGATGATG 1020
 Db 961 CCTAACCACTGACTTCTCTCTCTCTCCAGAAAGCTTCTTGTGTGATGATGATGATGATG 1020
 Qy 1021 CTGACCGGCGCATCTCTGCTTGTGAGGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080
 Db 1021 CTGACCGGCGCATCTCTGCTTGTGAGGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080
 Qy 1081 GTGGGCGCTTTTGTGATGACACCAAGGAGCATTTAGAGGCCCTTATCTTCAAGGCACT 1140
 Db 1081 GTGGGCGCTTTTGTGATGACACCAAGGAGCATTTAGAGGCCCTTATCTTCAAGGCACT 1140
 Qy 1141 ACCGCCCAACATCACTGCAATGGAATATCAAGGTGCCCAACCAACCGGAAACGTAAGG 1200
 Db 1141 ACCGCCCAACATCACTGCAATGGAATATCAAGGTGCCCAACCAACCGGAAACGTAAGG 1200
 Qy 1201 TGGGCTTGAATCTTCTTATGCTGTGAGCCCAAGCTACAGTGGGCTCTTGCACCAAG 1260
 Db 1201 TGGGCTTGAATCTTCTTATGCTGTGAGCCCAAGCTACAGTGGGCTCTTGCACCAAG 1260
 Qy 1261 ACTATGTGAAGATCAACCGGAGAAAGTACTGCGGTGAGAGGTCCAGTTTGTGTAGAGA 1320
 Db 1261 ACTATGTGAAGATCAACCGGAGAAAGTACTGCGGTGAGAGGTCCAGTTTGTGTAGAGA 1320
 Qy 1321 GCAACAGCAGCAAGATTAAGTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
 Db 1321 GCAACAGCAGCAAGATTAAGTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
 Qy 1381 TCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440

Db 1381 TCCTAGCTGAGTACCTCTCTCAGACCTCAACGACCCGCTGCCAGGATGTTACATGTGCA 1440
 QY 1441 AGACTGAGACGGTGATCTCCGAAAGAACTGCGCTGCGACGCGCTGGGAGACTGCCCGGATT 1500
 Db 1441 AGACTGAGACGGTGATCTCCGAAAGAACTGCGCTGCGACGCGCTGGGAGACTGCCCGGATT 1500
 QY 1501 ATAGTATGAGCGCTTACCTGCGATGCAATGCCACCCACAGTTTACAGTGCAGAAACCACT 1560
 Db 1501 ATAGTATGAGCGCTTACCTGCGATGCAATGCCACCCACAGTTTACAGTGCAGAAACCACT 1560
 QY 1561 TCTGCAAGCCCTCTTCTGCGGTCTGTGACAGTGTCAACGACTGTGGGAGCGAAGTAGCG 1620
 Db 1561 TCTGCAAGCCCTCTTCTGCGGTCTGTGACAGTGTCAACGACTGTGGGAGCGAAGTAGCG 1620
 QY 1621 AGAGGCGCTGACAGTGTCTGCTGGAGATTCAAGTGTTCGAAATGGGAAGTGTCTCCCTC 1680
 Db 1621 AGAGGCGCTGACAGTGTCTGCTGGAGATTCAAGTGTTCGAAATGGGAAGTGTCTCCCTC 1680
 QY 1681 AGAGCCAGAGTGTATGAGGAGAGACACTGTGAGATGGGTCTGACGAGGCTTCAATGTG 1740
 Db 1681 AGAGCCAGAGTGTATGAGGAGAGACACTGTGAGATGGGTCTGACGAGGCTTCAATGTG 1740
 QY 1741 ACAGCGTGAATGTCTCTCTTGCACCAATATACCTACCGCTGCCAAATGCGCTCTGTC 1800
 Db 1741 ACAGCGTGAATGTCTCTCTTGCACCAATATACCTACCGCTGCCAAATGCGCTCTGTC 1800
 QY 1801 TGAGCAAGGCAACCTCTGAGTGTGATGGAGAGCGAGCTGTAGCGATGGCTCCGATGAG 1860
 Db 1801 TGAGCAAGGCAACCTCTGAGTGTGATGGAGAGCGAGCTGTAGCGATGGCTCCGATGAG 1860
 QY 1861 AAAACGTGACTGTGGGCTGGCATCTCTTACCAAAAGAGCTGCGGTGTGGTGACGCA 1920
 Db 1861 AAAACGTGACTGTGGGCTGGCATCTCTTACCAAAAGAGCTGCGGTGTGGTGACGCA 1920
 QY 1921 ATGCGGACGAGGCGAGTGGCCCTGCGAGAGTGAAGCTCCACGCGCTGGGCGAGGCGACT 1980
 Db 1921 ATGCGGACGAGGCGAGTGGCCCTGCGAGAGTGAAGCTCCACGCGCTGGGCGAGGCGACT 1980
 QY 1981 TGTGTGGGCGCTGCGTCAATCTCTCTGACTGTGGTGTCTGTGAGCTCATTTGCTTTCAG 2040
 Db 1981 TGTGTGGGCGCTGCGTCAATCTCTCTGACTGTGGTGTCTGTGAGCTCATTTGCTTTCAG 2040
 QY 2041 ATGACAAAATTTCAATGATCTAGACTACAGATGTGAGCGGCTTCTGCGGTCTGCTCG 2100
 Db 2041 ATGACAAAATTTCAATGATCTAGACTACAGATGTGAGCGGCTTCTGCGGTCTGCTCG 2100
 QY 2101 ACCAGAGCAAGCGAGTGGCTCTGCGGAGTCAAGAGCTGAAGCTCAAAAGCTATCATCACC 2160
 Db 2101 ACCAGAGCAAGCGAGTGGCTCTGCGGAGTCAAGAGCTGAAGCTCAAAAGCTATCATCACC 2160
 QY 2161 ACCCTTCTCTTCAATGATTTCACTTGTGACTATGACATGCGCTTGTGAGCTGAGAAAT 2220
 Db 2161 ACCCTTCTCTTCAATGATTTCACTTGTGACTATGACATGCGCTTGTGAGCTGAGAAAT 2220
 QY 2221 CGGTGAGTACAGCAACCGTGTGGCGCCCATCTGCTGCTGATGATGTAACCTATGCTTTC 2280
 Db 2221 CGGTGAGTACAGCAACCGTGTGGCGCCCATCTGCTGCTGATGATGTAACCTATGCTTTC 2280
 QY 2281 CTGCTGAGCAAGCGCTGAGTCAAGAGCTGGGAGCACAAAAGAGGAGTACCGGAG 2340
 Db 2281 CTGCTGAGCAAGCGCTGAGTCAAGAGCTGGGAGCACAAAAGAGGAGTACCGGAG 2340
 QY 2341 CGCTGATCTCTGAGAAAGGCTGATGCTGCTGATCAACACGACCACTGTGAGACTCA 2400
 Db 2341 CGCTGATCTCTGAGAAAGGCTGATGCTGCTGATCAACACGACCACTGTGAGACTCA 2400
 QY 2401 TGCCGAGAGATACCCCAAGATGATGTGTGGTCTCTCAGTGGGGGTGTGAGACT 2460
 Db 2401 TGCCGAGAGATACCCCAAGATGATGTGTGGTCTCTCAGTGGGGGTGTGAGACT 2460
 QY 2461 CCTGCAAGGCTGACTCTGTGTGCGCCCTTGTCAAGGCGGAGAAAGATGGGCGAATGTTCC 2520
 Db 2461 CCTGCAAGGCTGACTCTGTGTGCGCCCTTGTCAAGGCGGAGAAAGATGGGCGAATGTTCC 2520

QY 2521 AGGCTGTGTGTGAGCTGGGAGTGAAGGCTGCGCTCAGAGAAACAAGCCAGGCGTGTACA 2580
 Db 2521 AGGCTGTGTGTGAGCTGGGAGTGAAGGCTGCGCTCAGAGAAACAAGCCAGGCGTGTACA 2580
 QY 2581 CAAGGCTCCCTGTATGTTTGGGAGCTGATCAAGAGCACTGGGATATATACAGATGAC 2640
 Db 2581 CAAGGCTCCCTGTATGTTTGGGAGCTGATCAAGAGCACTGGGATATATACAGATGAC 2640
 QY 2641 AGACAGCCGACCAAAACACCCACAGGATGCCCCGACATGACACCTGATACAGAGAG 2700
 Db 2641 AGACAGCCGACCAAAACACCCACAGGATGCCCCGACATGACACCTGATACAGAGAG 2700
 QY 2701 GAACACTGACGACATTTATCTGTGCGTCCCCCCCCCAACACAAACCAAGCTGTGACT 2760
 Db 2701 GAACACTGACGACATTTATGCTGTGCGTCCCCCCCCCAACACAAACCAAGCTGTGACT 2760
 QY 2761 GCATCTTATGAGACTCAGAGTCTTCCAAAGTGGGACCCCTCAAGAGTTGAGAGAGAAC 2820
 Db 2761 GCATCTTATGAGACTCAGAGTCTTCCAAAGTGGGACCCCTCAAGAGTTGAGAGAGAAC 2820
 QY 2821 TTGCGTGTAGCGGCCAGCTTGGGAGCAAGGCTTGTATGACAGCTTCCCTTACGCC 2880
 Db 2821 TTGCGTGTAGCGGCCAGCTTGGGAGCAAGGCTTGTATGACAGCTTCCCTTACGCC 2880
 QY 2881 CTGAGCTGGGTGAAGATGATGTCTGCTCCGAGAGCTGTCCAACTGTATGAGCTCCC 2940
 Db 2881 CTGAGCTGGGTGAAGATGATGTCTGCTCCGAGAGCTGTCCAACTGTATGAGCTCCC 2940
 QY 2941 GGGAGCCCTATGAGGAGGAGGCTCAGGCTCACTCTTTCAAGAAACGCGCAGCTTACGA 3000
 Db 2941 GGGAGCCCTATGAGGAGGAGGCTCAGGCTCACTCTTTCAAGAAACGCGCAGCTTACGA 3000
 QY 3001 ACCCGAGAAAGAGTGTACTTAAGGCTCAATTTGCTTGTGCTTCCAGGCGTGGAT 3060
 Db 3001 ACCCGAGAAAGAGTGTACTTAAGGCTCAATTTGCTTGTGCTTCCAGGCGTGGAT 3060
 QY 3061 TTGAGAGTAAACATTTTATTTCTTTTAAAAA 3106
 Db 3061 TTGAGAGTAAACATTTTATTTCTTTTAAAAA 3106

RESULT 2
 AAA88493
 ID AAA88493 standard; cDNA; 3149 BP.
 XX
 AC AAA88493;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Human matrixase cDNA.
 XX
 KW Matrilase; serine protease; human; breast cancer; pre-malignancy;
 KW actinic keratosis; leukoplakia; Barrett's epithelium;
 KW colunar metaplasia; ulcerative colitis; bowenoid papulosis;
 KW adenomatous colorectal polyp; Oyerat erythroplasia;
 KW vulvar intraepithelial neoplasia; tumour; metastasis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 36..2603
 FT /*tag= c
 XX
 PN W0200053232-A1.
 XX
 PD 14-SEP-2000.
 XX
 PF 10-MAR-2000; 2000MO-US006111.
 XX
 PR 12-MAR-1999; 99US-0124006P.
 XX
 PA (GBOU) UNIV GEORGETOWN.

QY	214	GCCCCAAGGCGCTGGGTGTGTCTGTGTGACGTGCTTTGACGCTTCTCTTGTCTCCCTCA	273
Db	188	GCCCCGGGCGCTGGGTGTGTGTCTGTGTGACGCGGTGTGTATGCGCTCTCTTGGTCTTGG	247
QY	274	TGCTGTGCTTGTGCTGTGTGTGCACTTCATATATGGAATGCGGGTTCAAAAGCTTCA	333
Db	248	GGATGCGCTTCTGTGTGTGSCATTTTGCATGCCGGAAGTGTCTGTCCAGAAAGTCTTCA	307
QY	334	ATGGCCATCTGAGGATCACAAATGAGATCTTCTGTGATGCGTATGAAATTCOACTCCA	393
Db	308	ATGGCTACATGAGGATCACAAATGAGAAATTTTGTGATGCTATCAGAACTCCACTCCA	367
QY	394	CAGAGTTTATCAGCTCTGCCACGACAGGTGAAGAGAGCGCTGAAGCTGCTGTCAATGAAG	453
Db	368	CTGAGTTTGAAGCTGTGGCAGCAAGGTGAAGACGGCTGAAGCTGTGTATACGGCGAG	427
QY	454	TCCCTGTCTGGGGTCCCTTACCAAGAAATGCGCTGTAACTGCGCTTCAGTGAAGGCAGTG	513
Db	428	TCCATTTCTGGGCCCTTACCAAGGAATGCGCTGTGAAGCGCTTCAAGAGGGGACAGG	487
QY	514	TCATGCGCTTACTGTGTGAGAGTTGAGATCCCCCACA CCTGSCAGAAAGATTGATC	573
Db	488	TCATGCGCTTACTGTGTGAGAGTTGAGATCCCGCAGCACTGTGTGAGAGGCCAGAC	547
QY	574	GCGCCATGTGCTGTGAGACGAGTTGTAACTTGGCACC CGGAGACGGGCACTGAAATCCT	633
Db	548	GCGTCATGTGGCAGAGACGGGTAGTCTAGTGTCCCGCGGGCGCGTCTCTTAATCTCT	607
QY	634	TGCTGTAAACA CTGTGTGTGGCTTCCCAATTGACCCCAATATGCTGCAGAGAGACTCAG	693
Db	608	TTGTGTACCTTCACTGTGTGTGCTTTCCCAAGCACTCCAAA CAGTACAGAGGACCCAG	667
QY	694	ACAAAGCTGCAATTTTGGCCCTGTGATCCCATGAGTGTGACAGTGA CAGCTTCACTACCC	753
Db	668	ACAAAGCTGCAAGCTTTGGCCTGTGACCGCGGGTGTGAGAGTGA TGCGCTTACACACCG	727
QY	754	CTGCTTTCCTCAACAGTCCCTTACCCCGGCGCATGCGCGCTGCGCAGTGGGTCTGCGGGGG	813
Db	728	CGGCTTCCCTGCAAGACCCCTTACCCCGCTCATGCGCGCTGCGCAGTGGGCTCTCGGGGG	787
QY	814	ACGGCACTCTGTGTGTGACCTCA CTTTCCGAACTTTGATGTGTGCTTCCCTGTGATGAGC	873
Db	788	ACGGCACTCACTGTGTGTGACCTCACTTTCGAGACTTTGACCTTGCCTCTGGAGGAGC	847
QY	874	ATGGCAGTGA CCTGTGTCA CCGGTGTATGATAGCTGAGCCCATGGAACCCACAGCTGTGG	933
Db	848	GCGGAGAGCACTGGTGA CCGGTGTATGAACAACCTTGAGCCCATGGAACCCACAGCTGTGG	907
QY	934	TGCGGCTGTGTGACCTTCTCACCCTCTCTCAACCTGACTTCTCTCTCTCCCAAGC	993
Db	908	TGCATGTTGTGTGGCACTTACCTCTCTCTCAACCTGACCTTCCACTCTCTCCAGAACG	967
QY	994	TCTTCTTGTCA CCGTGA TTAACCAATATGACCGGCGCATCTCGGCTTTGAGGCCACTT	1053
Db	968	TCTGTCTATCA CACTGA TTAACCAACTGAGCGGCGCATCCCGGCTTTGAGGCCACT	1027
QY	1054	TCTTCCAGCTGCCCCAAGATGAGCACTGTGGCGCTTTTGA GTGACACCCAGAGGACAT	1113
Db	1028	TCTTCCAGCTGCTCTAGATGAGCACTGTGAGGCGCGCTTACGTAAGCCACGAGGGACAT	1087
QY	1114	TTAGACGCCCCCTTACTATTCAGGCGCACTACCGGCGCAATCAATCAATGCAANTGAAATTC	1173
Db	1088	TCAACAGCCCCCTTACTACCGAGCGCACTACCA CCAATTGACTGTGCAATGGAACATTG	1147
QY	1174	AGGTGCCCCAACAACGGGAACGTGAAGGTGCGGCTTTCAAACTCTTCTATCTGTGTGACCCCA	1233
Db	1148	AGGTGCCCCAACAACGACATGTGAAGGTGCGGCTTTCAAACTCTTCTACTGTGTGAAGCCG	1207
QY	1234	ACGTACAGTGGGCTCTTGCACAAGACACTATGTGAGACTCAACGGGAGGAATCTGCG	1293
Db	1208	GCGTGTCTGTGGGCACTGTGCCCAAGGACTATGTGAGACTCAATGGGAGAAATCTGCG	1267
QY	1294	GTGAGAGCTCCAGTTTGTGTGTGAGCGACACGACGCAAGATTACAGTCCACTTCCATT	1353

Db	1268	GAGAGAGGTCCCAAGTTGCTGTCGACACAGAAACAGAAACAAGATCAAGTTCCGCTTCACT	1327
OY	1354	CTGATCACTGTTACACGAGACACCGGGTCTCTAGCTGAGTAACTCTCTACGATCCAAAG	1413
Db	1328	CAGATCACTCTTACACCGACACCGGCTCTTTAGCTGAATAACTCTCTACGATCCAAAG	1387
OY	1414	ACCGTGGCCACGGAGATGTTTCATGTCGCAAGACTGACCGGTGATCCGAAAGAACTGCGCT	1473
Db	1388	AACCATGCCCCGGGAGCATTTACAGTCCCGACCGGGCGGTGATACCGAAGAGCTCGCT	1447
OY	1474	GCGACGCTGGGACAGCTGCCGGATTAATAGTATGAGGGTTACTTCGCAATGGAATGCCA	1533
Db	1448	GTGATGCTGGGGCCCACTGCAACGACACAGGATAGCTCAACTCAAGTTGGACGCCG	1507
OY	1534	CCGACCACTTACCGTGAAGAAACAGATTTGCAAGGCCCTCTTCTGGGTCTGACAGTG	1593
Db	1508	GCCACCAATTCACGTGTCAGAAACAAAGTTTGCAGGCCCTCTTCTGGGTCTGACAGTG	1567
OY	1594	TCAACGACTGTGGGAGCGAAGTGAACGAGAGGGCTGCAGCTGTCTTGCTGGAGTTTCA	1653
Db	1568	TGAACGACTGCGGAGACAAACAGCAGACGAGGGGTGCAGTTGTCTCGGGCCCAACCTTCA	1627
OY	1654	AGTGTTCATATGGGAAGTGTCTCCCTCAGAGCAGAGGTAAATGGAAAGCAACTGTG	1713
Db	1628	GGTGTTCAATATGGGAAGTGTCTCTGAAAAGCCAGAGTGCATAATGGAAAGCACTGTG	1687
OY	1714	GAGATGAGGTCTGACGAGGCTTATGTGACAGCGCTGAATGTGTTCTTTCGACCAATAA	1773
Db	1688	GGGAGGGGTCCGACGAGGCTCTCGTCCCAAGGTGAACGTGCTCACTTGACCAAAACACA	1747
OY	1774	CCTACCGCTGCCAAATGGCCTTGTGCTGACGCAAGGGCAACCTGATGTGATGGGAGA	1833
Db	1748	CTTACCGCTGCTCAATGAGGCTCTGTGTTAAGCAAGGGCAACCTGATGTGACGGGAAG	1807
OY	1834	GGAGCTGTAGCGATGGCTCCGATGAGAAAACGTGACTGTGGGCTGCGATCCCTTTACA	1893
Db	1808	AGGACTGTAGCCAGCGCTCAGATGAGAAGGACTGCCACTGTGGGCTGCGATTCACGA	1867
OY	1894	AAACAGGCTGCGGTGTGTGTGTGACGAAATGCGGACGAGGGCGAGTGGCCCTTGATAGGTGA	1953
Db	1868	GACAGGCTGTGTGTGTGTGTGTGGGGGACGAGATGCGGATGAGGGCGAGTGGCCCTTGATAGGTGA	1927
OY	1954	GCCTCCAGCCCTGGGGCCAGGGGCACTGTGTGTGGGGCTGCGTCACTCTCTCTACTGAGC	2013
Db	1928	GCTGTATCTCTGGGSCAAGGGCAATCTGCGGTGCTTCCCTCATCTCTCCCAACTGCG	1987
OY	2014	TGGTCTCTCAGTCAATTTGCTTTCAAGATGACAAAATTTCAATACTCAGACTACAGA	2073
Db	1988	TGGTCTCTCAGCACTGTATCATGATGACAGAGATTCAAGTACTCAGACCCCAAGC	2047
OY	2074	TGTGACGCGCTTCTTGAGTCTGTGTGACCAAGACCAAGCCGAGCTCTGGGGGTGACAG	2133
Db	2048	AGTGAACGCGCTTCTTGAGGCTTTCACAGACACGAGCCGAGCGGCCCTTGAGGTGACAG	2107
OY	2134	AGTGAAGCTCAACGATCATCAACCAACCCCTTCTTCAATGATGATTTCACTTTCAGTAAG	2193
Db	2108	AGGCAAGGCTCAAGCCGATCATCTCCACCCCTTCTTCAATGATTTCACTTTCAGTAAG	2167
OY	2194	ACATGCGCTTGTGAGCTGAGAGAAAGTGGTGAATGACAGACCGTGTGGGCCCAACT	2253
Db	2168	ACATGCGCTGTGAGCTGAGAGAAACCGGACAGATGACAGCTCATAGTGTGGGCCCAACT	2227
OY	2254	GCTTGCTGATGTTGTTACCATGATCTTCCCTGCTGACAGGCAATCTGGGTTCACAGGCTGG	2313
Db	2228	GCTTGCTGAGCGCTCCCATGATCTTCCCTGCTGACAGGCAATCTGGGTTCACAGGCTGG	2287
OY	2314	GGCACACAAAGAGGAGGTATCCGAGACGCTGATCTTGACAGAGGGTGAATCCGTGTCA	2373
Db	2288	GACACACCGATATGAGAGCACTGGCGGCTGATCTTGCAAAAGGATGAGATCCCGTCA	2347
OY	2374	TCAACCAACCACTGTGAGGACTCATGCGGACAGATCAACCCCAGAATGATGTGTG	2433

Db 2348 TCAACGACACCTGCGAGACCTCTGCGGAGAGATCAGCCGCGATGATGTCG 2407
 QY 2434 TGGTTTCTCAGTGGGGGTGTGAGATCTCTGCGAGGAGTCTGTGTCCTTGTCA 2493
 Db 2408 TGGCTTCTCAGGCGGCGGTGAGATCTCTGCGAGGAGTCTGTGTCCTTGTCA 2467
 QY 2494 GCGCGAGAAAGATGGGCGAATGTTCCAGGCTGTGTGTGAGCTGGGGTGAAGCTGCG 2553
 Db 2468 GCGTGAAGGCGGATGGGCGGATCTTCCAGGCGGCTGTGTGAGCTGGGGAGAGCTGCG 2527
 QY 2554 CTGAGAGGAAAGAGGCGGCTGTGAGCAAGGCTCTGCTGTGAGTCTGGGAGTGTCAAG 2613
 Db 2528 CTGAGAGGAAAGAGGCGGCTGTGAGCAAGGCTCTGCTGTGAGTCTGGGAGTGTCAAG 2587
 QY 2614 AGCAGCTGGGCTGTGAGCAAGGCTGTGAGCAAGGCTCTGCTGTGAGTCTGGGAGTGTCAAG 2645
 Db 2588 AGAACAAGGCTGTGAGGCGGCGGAGCAGCCCA 2619

RESULT 4
 AAX87815
 ID AAX87815 standard; cDNA; 3147 BP.
 AC AAX87815;
 DT 09-NOV-1999 (first entry)
 DE Tumour antigen derived gene-15 (TADG-15) cDNA.
 KW Tumour antigen derived gene-15; TADG-15; serine protease; human;
 KM breast cancer; ovary cancer; carcinoma; diagnosis; ss.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 23..2590
 FT /*tag= a
 PN MO9942120-A1.
 PD 26-AUG-1999.
 PF 18-FEB-1999; 99WO-US003436.
 PR 20-FEB-1998; 98US-00027337.
 PA (UVAR-) UNIV ARKANSAS.
 PI O'Brien TJ, Tamimoto H;
 DR MPI: 1999-527418/44.
 DR P-PSDB; AAY06671.
 PS A new extracellular serine protease for diagnosis of neoplastic disease.
 XX Claim 2; Fig 9; 71pp; English.
 CC This is the nucleotide sequence of an isolated cDNA that codes for an
 CC extracellular serine protease, termed tumour antigen derived gene-15
 CC protein (see AAY06671), that is overexpressed in breast and ovarian
 CC carcinomas. The TADG-15 gene can be used as a diagnostic and therapeutic
 CC target in ovarian carcinoma and other carcinomas including breast,
 CC prostate, lung and colon. The TADG-15 cDNA was isolated from ovarian
 CC carcinoma by PCR using primers directed to conserved areas of the serine
 CC protease family. The invention also provides: a vector that is capable of
 CC expressing DNA encoding TADG-15 protein; host cells selected from
 CC bacterial cells (especially Escherichia coli), mammalian cells, plant
 CC cells and insect cells; and a method of detecting expression of TADG-15
 CC protein using a hybridisation probe
 SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Best Local Similarity 81.2%; Pred. No. 0;
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
 QY 45 GATGGACCGGCAAAACCAATGGGTAGCAATGGGGGCGGAGCGGAGGCGCTCTCAG 104
 Db 5 GAGCGGCTTCGGGGTACCAATGGGGAGGATGGGGCGGAGCGGAGGCGGAGGCGGAG 64
 QY 105 GACTTCGGGCGGAGCTCAAGTCAACTCCCGGCTAGAGAACTGATGGCTTTGAGAG 164
 Db 65 GACTTCGGGCGGAGCTCAAGTCAACTCCCGGCAAGAGAAATGATGGCTTTGAGAG 124
 QY 165 GGTGTGAAGTTCTGCTCGGAACTTCGCAAGAAAGTGAAGAGGAGGAGGAGGAGGAG 224
 Db 125 GGGTGAAGTTCTGCTCGGAACTTCGCAAGAAAGTGAAGAGGAGGAGGAGGAGGAG 184
 QY 225 TGGGT 284
 Db 185 TGGGT 244
 QY 285 CTGATGTGCACTTCATATGCAATGCGGAGTTCAAAGAGTTCATGAGGAGTTCG 344
 Db 245 CTGATGTGCACTTCATATGCAATGCGGAGTTCGATGCAAGAGTTCATGAGGAGTTCG 304
 QY 345 AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTTCACTCAAGATTTATC 404
 Db 305 AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTTCACTCAAGATTTATC 364
 QY 405 AGCTGGCGGAGCGAGTGAAGAGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 464
 Db 365 AGCTGGCGGAGCGAGTGAAGAGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 424
 QY 465 GGTGCTTACCAAGAGT 524
 Db 425 GGTGCTTACCAAGAGT 484
 QY 525 TACTGTGAGATGAGATCTTCCCACTGCGAGAAAGAGTGTGTGTGTGTGTGTGTGT 584
 Db 485 TACTGTGAGATGAGATCTTCCCACTGCGAGAAAGAGTGTGTGTGTGTGTGTGTGT 544
 QY 585 GTGAGCGAGTGTGAATGCGACCGGAGAGCGGAGCTGTAATCTCTGTGTGTGTGT 644
 Db 545 GAGGAGCGGATGATGATGCTGCGGAGGAGCGGAGCTGTAATCTCTGTGTGTGTGT 604
 QY 645 TCTGTGTGAGCTTCCCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 704
 Db 605 TCAGT 664
 QY 705 AGTTTGT 764
 Db 665 AGTTTGT 724
 QY 765 AACAGTCCCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824
 Db 725 GAGAGCGGATGATGATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 784
 QY 825 GTGCTGAGCTTACCTTCCAGAGCTTTGATGTGCTCTCTGTGATGAGCTGAGCTGAG 884
 Db 785 GTGCTGAGCTTACCTTCCAGAGCTTTGATGTGCTCTCTGTGATGAGCTGAGCTGAG 844
 QY 885 CTGCTACCGTGTATGATGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 944
 Db 845 CTGCTACCGTGTATGATGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 904
 QY 945 GAGAGCTTCTACCGTGTATGAGCTTCTCTCTGTGATGAGCTGAGCTGAGCTGAG 1004
 Db 905 GAGAGCTTCTACCGTGTATGAGCTTCTCTCTGTGATGAGCTGAGCTGAGCTGAG 964
 QY 1005 AGCTGTATTAACCAATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1064
 Db 965 ACAGTGTATTAACCAATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1024
 QY 1065 CCCAAGATGAGAGCTGT 1124

Db 1025 CCTAGATGAGCAGCTGTGAGGCGCTTACGTAAGCCGAGGACATTCAACAGCCCC 1084
 QY 1125 TACTATCCAGGCGCACTACCCGCCAACATCACTGCACATGGAATATCAAGGTGCCAAC 1184
 Db 1085 TACTACCCAGGCGCACTACCCAACCAATGACTGCACATGGAACATTTAGGGGCCAAC 1144
 QY 1185 AACCGGAAAGTGAAGGTGGGCTTCAAACTCTTATCTGGTGAACCCCAAGTACCAGT 1244
 Db 1145 AACCAACATGTGAAGGTGAAGCTTCAAAATTTCTTACTGCTGAGGCCGGGCTGTGG 1204
 QY 1245 GGCTCTGCACCAAGGACTATGTGAGATCAACGAGGAGAAATCTCCGGTGAAGGCTC 1304
 Db 1205 GGGACCTGCCCCGAAGACTACGTGAGATCAATGGGGAGAAATACGCGGAGAGAGCTCC 1264
 QY 1305 CAGTTTGTGTGAGCAGCAACAGACAGATTAACGTCACATTTCCATTCTATCTCTG 1364
 Db 1265 CAGTTGTGTGAGCAGCAACAGATCAAGATCAAGTTGCTTCCACTCACTAGATCAGTCC 1324
 QY 1365 TACACGGACACGAGGCTTCTAGGCTAGTACCTCTCCAGACTCCAAAGACCCGAGCCCA 1424
 Db 1325 TACACGGACACGAGGCTTCTAGGCTAGTACCTCTCCAGACTCCAGTACCCATGCCCC 1384
 QY 1425 GGGATGTTCAATGTGCAAGACTGAGCAGTGCATCCGAAAGGAACTGCGCTGCGAGCTGG 1484
 Db 1385 GGGCAGTTCAAGTCCGCAACGAGGCGGTGTATCCGGAAGAGCTGCGCTGTGATGGCTGG 1444
 QY 1485 GCAGACTGCCGGAATTAATGATGAGCGTTACTGCCGATGCAATGCAACCCACCACTTC 1544
 Db 1445 GCCGACTGCACCCACACACAGATGAGCTCAACTGCAAGTTGCGAGCGCGGCGCAACTTC 1504
 QY 1545 ACGTGCAAAACCAAGTTCTGCAAGCCCTCTCTCGGGTCTGTGACAGTGTCAACGACTGT 1604
 Db 1505 ACGTGCAAAACCAAGTTCTGCAAGCCCTCTCTCGGGTCTGTGACAGTGTGAAAGACTGC 1564
 QY 1605 GGGGACGGAAGTGCAGAGAGGAGCTGCAGCTGTCTCTGAGGATTTCAAGTGTCCAAT 1664
 Db 1565 GGAAGCAACAGCGACGAGAGGAGGTGCAAGTTGTCCGGCCAGACCTTCAGGTGTTCCAAT 1624
 QY 1665 GGGAGAGTGTCTCCCTCAGAGCCAGAGTGTATGGAAGAGCAACTGTGAGATGGGTCT 1724
 Db 1625 GGGAGAGTGTCTCCCTCAGAGCCAGAGTGTGCAAGTGTGAGAGAGCGACTGTGGGAGG 1684
 QY 1725 GAGAGGCTTCATGTGACAGCGGTGATGTGTCTCTTGACCAAAATATTAACCTACCGCTGC 1784
 Db 1685 GAGAGGCTTCATGTGACAGCGGTGATGTGTCTCTTGACCAAAATATTAACCTACCGCTGC 1744
 QY 1785 CAAAATGAGCCTCTGTCTGAGCAAGGCAACCCCTGATGTGTATGGAAGAGCGAGCTGAGC 1844
 Db 1745 CTCAATGGGCTCTCTGTGAGCAAGGCAACCCCTGATGTGTATGGAAGAGCGAGCTGAGC 1804
 QY 1845 GATGGCTCCGATGGAAGAACTGTGACTGTGGGCTGTGCAATCTTTTACCAGAGCTGCGC 1904
 Db 1805 GACGGCTCAGATGGAAGAGTGTGACTGTGGGCTGTGCAATCTTTTACCAGAGCTGCGC 1864
 QY 1905 GTGGTGGTGGCAGGAATGCGGACGAGGGCGAGTGGCCCTGAGTGAAGCTTCACAGCC 1964
 Db 1865 GTTGTGGGGGCAAGATGCGGATGAGGGCGAGTGGCCCTGAGTGAAGCTTCAGATGCT 1924
 QY 1965 CTGGGCGCAGGGCCAGATCTGGGGTGTCTCCCTCATCTCTCCCACTGGGCTGTCTGGCC 1984
 Db 1925 CTGGGCGCAGGGCCAGATCTGGGGTGTCTCCCTCATCTCTCCCACTGGGCTGTCTGGCC 1984
 QY 2025 GCTCATTTGCTTTCAGAGATGACAAAAATTTCAAGTACTCAGACTACAGATGTGAGCGGCG 2084
 Db 1985 GCACACTGCTCATCTGATGACAGAGATTCAAGTACTCAGACTACAGATGTGAGCGGCG 2044
 QY 2085 TTCTGGGCTCTGGAGCAGAGCAAGCGCAGTGGCTCTGGGGTGCAGAGCTGAAGCTC 2144
 Db 2045 TTCTGGGCTCTGGAGCAGAGCAAGCGCAGTGGCTCTGGGGTGCAGAGCTGAAGCTC 2104
 QY 2145 AAAGCATATCACCACCTCTCTCAATGATTTACCTTGCATGACTGACATGCGCTTG 2204
 Db 2105 AAAGCATATCACCACCTCTCTCAATGATTTACCTTGCATGACTGACATGCGCTTG 2164

QY 2205 CTGAGCTGAGAGATGCGTGTGAGTACAGACCGCTGTGTCGCCCCATCTGCTGCTGAT 2264
 Db 2165 CTGAGCTGAGAGATGCGTGTGAGTACAGACCGCTGTGTCGCCCCATCTGCTGCTGAT 2224
 QY 2265 GCTAACCATGTCTTCCCTGTGCGCAAGGCCATCTGGGTCAAGAGCTGGGGGCAACAAAA 2324
 Db 2225 GCTTCCATGTCTTCCCTGTGCGCAAGGCCATCTGGGTCAAGAGCTGGGGGCAACAAAA 2284
 QY 2325 GAGGAGGTACCGGAGCGCTGATCTCTGCAAGAGGTGAGATCCGTGTATCAACAGACC 2384
 Db 2285 TATGAGGCACTGGCGCGCTGATCTGCAAAAGGTGAGATCCGTGTATCAACAGACC 2344
 QY 2385 ACCTGTGAGAGACTCATGCGCGCAGCAGATCACCCCAAGATGATGTGTGTGTTCTTC 2444
 Db 2345 ACCTGTGAGAGACTCATGCGCGCAGCAGATCACCGCGCATGATGTGTGTGTTCTTC 2404
 QY 2445 AGTGGGGTGTGAGCTCTGCGCAGGCTGACTGTGTGTGCTTGTTCAGAGCGGAGAA 2504
 Db 2405 AGCGCGGGGTGAGCTCTGCGCAGGCTGATTCGAGGGGACCCCTGTCCAGCGTGAAGCG 2464
 QY 2505 GATGGCGAATGTTCCAGGCTGTGTGTGAGCTGGGGGTGAAGGCTGCGCTCAGAGAAC 2564
 Db 2465 GATGGCGAATGTTCCAGGCTGTGTGTGAGCTGGGGGTGAAGGCTGCGCTCAGAGAAC 2524
 QY 2565 AAGCCAGCGGTGTACACAAAGGCTCCCTGTATGTGAGCTGATCAAAAGACACTGGG 2624
 Db 2525 AAGCCAGCGGTGTACACAAAGGCTCCCTGTATGTGAGCTGATCAAAAGACACTGGG 2584
 QY 2625 GTATAGCAGATGACACAGACAGCCGACCAACACCCACAGAGGATCCGACATGAGCA 2684
 Db 2585 GTATAGGAGCGCGG---GCCACCCAAATGTGTACACTGTGGGGCCACCATGTGTCCAC 2641
 QY 2685 CCTGATACAGAGAGGAACTGTACAGCAATTATGTGTGGGCTCCCGCCCAACACA 2744
 Db 2642 CAGGTGTGACG-CCTGACAGCTGTGAGACTGTGACCGCTGACTGACACAGGCGC-CCAGA 2699
 QY 2745 ACCCAGACTGTGAACCTGATCCTTAGACTCAAGT 2780
 Db 2700 ACATACACTGTGAACCTCAATCTCCAGGGCTCCAAAT 2735

RESULT 5
 AAH23609/c
 ID AAH23609 standard; RNA; 3147 BP.
 XX
 AC AAH23609;
 XX
 DT 03-AUG-2001 (first entry)
 XX
 DE Human TADG-15 antisense RNA sequence.
 XX
 KM TADG-15; cytosolic; vaccine; ovarian tumour; cancer; human; antisense;
 KM tumour antigen-derived gene 15; serine protease; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200129056-A1.
 XX
 XX 26-APR-2001.
 PD
 PF 20-OCT-2000; 2000MO-US029095.
 XX
 PR 20-OCT-1999; 99US-00421213.
 XX
 PA (UTAR-) UNIV ARKANSAS.
 XX
 PI O'Brien TV, Tanimoto H;
 XX
 DR WPI, 2001-381031/40.
 XX
 PT Novel extracellular serine protease, termed tumor antigen-derived gene 15 protein overexpressed in carcinomas and DNA encoding it, for diagnosis,

XX treatment, prevention of cancer, particularly breast, ovarian cancer.
XX
XX Example 8; Page 102-103; 130pp; English.

CC The present invention relates to human tumour antigen-derived gene 15
CC (TADG-15) protein and coding sequence (see AAH213601 and AB98500). TADG-
CC 15 is an extracellular serine protease. It was found that TADG-15 is over
CC expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20
CC residues that lack TADG-15 protease activity are useful for vaccinating
CC an individual against TADG-15, having, suspected of having or at risk of
CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic
CC or therapeutic target in cancer. TADG-15 was cloned and expressed in the
CC opposite orientation such that an antisense RNA molecule (the present
CC sequence) was produced. The present sequence is useful for hybridising to
CC the complementary RNA in a cell and thereby inhibiting translation of
CC TADG-15 into protein

XX Sequence 3147 BP; 583 A; 958 C; 952 G; 0 T; 654 U; 0 Other;

Query Match 60.6%; Score 1883.2; DB 4; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATCGACCGCCAAAACCATGGGTAGCATCGGGCCGGAGAGCCGGAGGGGCTCTG 104
DB 3143 GAGCGGCGCTCGGGGTACCATGGGAGCCATCGGACCCGCAAGGCGAGGGGCCCGGAAG 3084
QY 105 GACTTCGGCGGGGAGCTCAAGTAACTCCCGCTAGAGACATGATGGCTTTGAGAG 164
DB 3083 GACTTCGGCGGGGAGCTCAAGTAACTCCCGCGACGAGAAAGTGAATGGCTTTGAGAGAA 3024
QY 165 GGTGGAGATTCCTCTGCTGCGCAATGCGAAGAAAGTGGAGAGCGAGGCCCGCAGGCGC 224
DB 3023 GCGTGGAGATTCCTCTGCTGCGCAATGCGAAGAAAGTGGAGAGCGAGGCCCGCAGGCGC 2964
QY 225 TGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 284
DB 2963 TGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2904
QY 285 CTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 344
DB 2903 CTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2844
QY 345 AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTCCACTCCACAGAGTTTATC 404
DB 2843 AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTCCACTCCACAGAGTTTATC 2784
QY 405 AGCCTGGCAGCAGGTGAGAGAGGCGTGAAGCTGCTGTCAATGAGTCCCTGTCTG 464
DB 2783 AGCCTGGCAGCAGGTGAGAGAGGCGTGAAGCTGCTGTCAATGAGTCCCTGTCTG 2724
QY 465 GGTCTCTACCAAGAAAGTGGCTGTAACTGCTTCACTGAGGGGAGTGTACGCTAC 524
DB 2723 GGGCCCTACCAAGAAAGTGGCTGTAACTGCTTCACTGAGGGGAGTGTACGCTAC 2664
QY 525 TACTGTGAGGTTCAGATCCCGCCCAACTGCGAGAAAGTTGATGCGCGCATGCT 584
DB 2663 TACTGTGAGGTTCAGATCCCGCCCAACTGCGAGAAAGTTGATGCGCGCATGCT 2604
QY 585 GTGAGAGAGTTGTAACATGTCACCCCGAGACGCGGACAGAAATCCTTGCTGTAA 644
DB 2603 GAGAGAGCGGTGTAACATGTCACCCCGAGAGCGGCGGCTCTGTAAGTCTTTGTGTCAC 2544
QY 645 TCTGTGTGCTTCTTCCCATTTGACCCCAAGATGCTGACAGAGACTCAGACAGAGCTGC 704
DB 2543 TCACTGTGTGCTTCTTCCCATTTGACCCCAAGATGCTGACAGAGACTCAGACAGAGCTGC 2484
QY 705 AGTTTGGCTGAGTGCATGCTGAGAGAGTGCAGAGTGCAGCTTCACTACCTCTGCTTCCC 764
DB 2483 AGCTTTGGCTGAGTGCATGCTGAGAGAGTGCAGAGTGCAGCTTCACTACCTCTGCTTCCC 2424
QY 765 AACAGTCCCTAACCGGCGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824

DB 2423 GACAGCCCTTACCCCGCTCATGCCCGCTGCACTGAGGCGCTCGGGGGAGCGCCGACTCA 2364
QY 825 GTGCTGAGCTCACTTCCGAGGCTTGTGATGCTCCCTGTGATAGCATGGCACTGAC 884
DB 2363 GTGCTGAGCTCACTTCCGAGGCTTGTGATGCTCCCTGTGATAGCATGGCACTGAC 2304
QY 885 CTGGTACCGGTATGATGATGAGCCCGCATAGAAACCCGACGCTGAGGCGGCTGTG 944
DB 2303 CTGGTACCGGTATGATGATGAGCCCGCATAGAAACCCGACGCTGAGGCGGCTGTG 2244
QY 945 GGCACCTTTCACTCCCTCTCAACCTGACTTCTCTCTCCGAGAGCTTCTCTGTC 1004
DB 2243 GGCACCTTTCACTCCCTCTCAACCTGACTTCTCTCTCCGAGAGCTTCTCTGTC 2184
QY 1005 AGCGTATTAACCAATCTGACCCGGGACATCTCTGCTTTGAGGCCACTTTTCAAGCTG 1064
DB 2183 AACCTGATTAACCAATCTGACCCGGGACATCTCTGCTTTGAGGCCACTTTTCAAGCTG 2124
QY 1065 CCCAAGATGAGCAGCTGTGCGGCTTTTGTGAGTACACCCAGGACATTTGACAGCCCG 1124
DB 2123 CCTAGGATGAGCAGCTGTGAGGCGGCTTACGTAAAGCCAGGACATTTCAACAGCCCG 2064
QY 1125 TACTATCCAGGCGCTACCTCCGCTCAACATCACTGCACTGCAATGCAATTCAGGTGCCAAC 1184
DB 2063 TACTATCCAGGCGCTACCTCCGCTCAACATTCAGTGCATGGAACATGAGGTGCCAAC 2004
QY 1185 AACCGGAACGTGAAGGTGCGTTCAACTCTTCTATCTGTGTGAGACCCCAAGTACCAATG 1244
DB 2003 AACCGGAACGTGAAGGTGCGTTCAACTCTTCTATCTGTGTGAGACCCCAAGTACCAATG 1944
QY 1245 GGCCTCTGACCAAGAGCTATGTGAGATCAACGGGGAGAAATGACTCGGTGAGAGGTCC 1304
DB 1943 GGCACCTGCGCCCAAGAGCTATGTGAGATCAATGGGAGAAATTCGCGGAGAGAGGTCC 1884
QY 1305 CAGTTGTGTGAGAGCAGACAGACAGCAAGTTACATCACTTCATCTGATCACTG 1364
DB 1883 CAGTTGTGTGAGAGCAGACAGACAGCAAGTTACATCACTTCATCTGATCACTG 1824
QY 1365 TACACGAGACCGGGTCTCTGAGTAACTCTCTCACTGAGTCCCAAGAGCCGTCCTCA 1424
DB 1823 TACACGAGACCGGGTCTCTGAGTAACTCTCTCACTGAGTCCCAAGAGCCGTCCTCA 1764
QY 1425 GGGATGTCATGTGACAGACTGACGCTGATCCGAAAGAACTGCGTGCAGCGGTGG 1484
DB 1763 GGGAGATTCAGGTGCGGACCGGGGCGGTATCCGGAAGAGCTCGCTGTGATGCTGG 1704
QY 1485 GCAATCTCCCGGATTAATGATGAGGCTTACTGCGATGCAATGCCACCAAGTTC 1544
DB 1703 GCGCATGACACCGAACACAGAGATGAGTCAACTGACATGTGAGCGCGGCCACCAAGTTC 1644
QY 1545 ACGTGCAAAACCAAGTCTGCAAGCCCTCTTCTGGGCTGTGACAGTGTCAACGACTGT 1604
DB 1643 ACGTGCAAAACCAAGTCTGCAAGCCCTCTTCTGGGCTGTGACAGTGTCAACGACTGT 1584
QY 1605 GGGAGCGAAGTGAAGAGAGGCTGACGCTGCTCTGCGGAGGTTTCAAGTGTCCAT 1664
DB 1583 GGAAGCAACAGCGACAGAGAGGCTGACGCTGCTCTGCGGAGGTTTCAAGTGTCCAT 1524
QY 1665 GGGAGAGTGTCTCTCTCAGAGCAGAAAGTATGAGAAAGCAACTGTGAGAGTGGTCT 1724
DB 1523 GGGAGAGTGTCTCTCTCAGAAAGCAGAGTGAATGGGAAAGGACGCTGTGGGAGCGGTCTC 1464
QY 1725 GACGAGGCTTCACTGACAGGCTGAATGCTCTCTGACCAAAATTAATTAACCTACCGCTGC 1784
DB 1463 GACGAGGCTTCTCTCTCAGAGGCTGAATGCTCTCTGACCAAAATTAATTAACCTACCGCTGC 1404
QY 1785 CAAATGAGCTCTCTCTGAGCAGAGGCAACCTGAGTGTGAGAGAGAGAGAGCTGAGC 1844
DB 1403 CTCAATGAGGCTCTCTCTGAGAGAGGCAACCTGAGTGTGAGAGAGAGAGAGCTGAGC 1344
QY 1845 GATGCTCTCCATGAGAAACCTGATGCTGTGGCTGTGATCTTTACCAACAGGCTGC 1904
DB 1343 GACGCTCAATGAGAAAGCTGAGCTGTGGCTGTGATCTTTACCAACAGGCTGC 1284

```

QY 1905 GTGTTGTGTCGACGAAATGCGGACGAGGCGGAGTGGCCCTTGGCAGGTGAGCTTCCAGCGC 1964
XX |||
OS |||
Db 1283 GTTGTGGGGGACGATGCGGATGAGGCGGAGTGGCCCTTGGCAGGTGAGCTTCCAGCT 1224
QY 1965 CTGGGCGACGAGGCGCACTGTGTGGGGGCTGGCTCATCTCTCTGACCTGGCTGTCTGTGA 2024
XX |||
FT |||
Db 1223 CTGGGCGACGAGGCGCACTGTGTGGGGGCTGGCTCATCTCTCTGACCTGGCTGTCTGTGC 1164
QY 2025 GCTCATGCTTTTCAGGATGACAAAATTTCAAGTACTGAGACTACAGATGTGGACGCGC 2084
XX |||
OS |||
Db 1163 GCACACTGTCTACATGATGACAGAGATTGAGTACTCAGACCCCGCAGAGTGGACGCGC 1104
QY 2085 TTCTCTGGGTCTGTGACACGAGCAAGCGGAGTGGCTCTGGGGTGGAGGACGTGAAGCTC 2144
XX |||
FT |||
Db 1103 TTCTCTGGGTCTGTGACACGAGCAAGCGGAGTGGCTCTGGGGTGGAGGACGTGAAGCTC 1044
QY 2145 AAACGTATCATCACCACCTTCTTCAATGATTTTCACTTTCAGCTATGACATGCGCTTG 2204
XX |||
OS |||
Db 1043 AAGCGCATCATCTCCACACCTTCTTCAATGATTTTCACTTTCAGCTATGACATGCGCTTG 984
QY 2205 CTGGAGCTGAGAAAGTGTGTGAGTACAGCAGCTGTGGCCCATCTGCTGCTGAT 2264
XX |||
OS |||
Db 983 CTGGAGCTGAGAAAGTGTGTGAGTACAGCAGCTGTGGCCCATCTGCTGCTGAT 924
QY 2265 GCTACCCATCTCTTCCCTGCTGTGGCAAGGCGCATGCGGCTCAAGGCTGGGGGCGACACAAA 2324
XX |||
OS |||
Db 923 GCTCTCCATGTCTTCTCTGCGGCAAGGCGCATGCGGCTCAAGGCTGGGGGCGACACAAA 864
QY 2325 GAGGAGGTACCCGAGCGCTGATCTCTGACAGAGGCTGAGTCCGTGTATCAACCCAGACC 2384
XX |||
OS |||
Db 863 TATGGAGGACCTGGCGCGTGTATCTTCCAAAAGGCTGAGTCCGCGTATCAACCCAGACC 804
QY 2385 ACCGTGAGGACCTCATGCTCGGACGAGATCAACCCAGAAATGATGTGTGGGTTTCTC 2444
XX |||
OS |||
Db 803 ACTGCGAGAACTCTCTGCGGACGAGATCAACCCAGAAATGATGTGTGGGTTTCTC 744
QY 2445 AGTGGGGTGTGACATCTCTGCGGAGTGTGAGTCTGTGTGGCCCTTGTCAAGGCGGAGAAA 2504
XX |||
OS |||
Db 743 AGGCGCGGCTGTGACATCTCTGCGGAGTGTGAGTCTGTGTGGCCCTTGTCAAGGCGGAG 684
QY 2505 GATGGGCGAAATGTTCCAGGCTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2564
XX |||
OS |||
Db 683 GATGGGCGAAATGTTCCAGGCTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 624
QY 2565 AAGCCAGGCTGTGACCAAGGCTCTCTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2624
XX |||
OS |||
Db 623 AAGCCAGGCTGTGACCAAGGCTCTCTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 564
QY 2625 GTATAGCAGATGAGACAGACGACGACCAACCAACCAAGGATGGCGGACATGACA 2684
XX |||
OS |||
Db 563 GTATAGGAGGCGGAGG--GCAACCAATGTGTACCTCTGGGGGCGACCACTGTCCACC 507
QY 2685 CCTGGATACAGAGAGGAACTGACGACATTTATGCTGTGGCTCTCCCCCGCAACACA 2744
XX |||
OS |||
Db 506 CCAAGTGTGACAG--CCTGACGCTGTGAAGCTGACCGCTGACTGACACGACGCGCC--CCAGA 449
QY 2745 ACCCAAGCTGTGAAGTGTGATCTTTAGACTGAGACT 2780
XX |||
OS |||
Db 448 ACATACACTGTGAACCTCAATCTCCAGGGGCTCAAAAT 413

```

```

RESULT 6
AAH23601
ID AAH23601 standard; DNA; 3147 BP.
XX AC
XX AAH23601;
XX 03-AUG-2001 (first entry)
XX DE Human TADG-15 coding sequence.
XX KW Human; TADG-15; cytoskeletal; vaccine; ovarian tumour; cancer;

```

```

KW tumour antigen-derived gene 15; extracellular serine protease; ds.
XX Homo sapiens.
OS Key Location/Qualifiers
XX CDS 23..2590
XX FT /*tag=a
XX FT /product="Human TADG-15"
XX PN MO200129056-A1.
XX PD 26-APR-2001.
XX PF 20-OCT-2000; 2000MO-US029095.
XX PR 20-OCT-1999; 99US-00421213.
XX (UYAR-) UNIV ARKANSAS.
XX O'brien TJ, Tanimoto H;
XX PI
XX DR MPI: 2001-381031/40.
XX DR P-PSDB; AAB98500.
XX PT Novel extracellular serine protease, termed tumor antigen-derived gene 15
XX PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,
XX PT treatment, prevention of cancer, particularly breast, ovarian cancer.
XX PS Claim 2; Fig 2; 130pp; English.
XX CC The present sequence is the coding sequence for human tumour antigen-
XX CC derived gene 15 (TADG-15) protein. TADG-15 is an extracellular serine
XX CC protease. It was found that TADG-15 is over-expressed in ovarian tumors.
XX CC TADG-15 protein or its fragments of 9-20 residues that lack TADG-15
XX CC protease activity are useful for vaccinating an individual against TADG-
XX CC 15, having, suspected of having or at risk of getting cancer.
XX CC Furthermore, the TADG-15 gene can be used as a diagnostic or therapeutic
XX CC target in cancer
XX CC
XX SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

```

```

Query Match 60.6%; Score 1883.2; DB 4; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

```

```

QY 45 GATCGGACCGCCAAACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 104
XX |||
Db 5 GAGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 64
QY 105 GACTTCGGCGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 164
XX |||
Db 65 GACTTCGGCGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 124
QY 165 GCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 224
XX |||
Db 125 GCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 184
QY 225 TGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 284
XX |||
Db 185 TGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 244
QY 285 CTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 344
XX |||
Db 245 CTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 304
QY 345 AGATTCACAAATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 404
XX |||
Db 305 AGATTCACAAATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 364
QY 405 AGCTGGCGCAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 464
XX |||
Db 365 AGCTGGCGCAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 424

```

QY 465 GGTCCCTACCAAGAGTGGCTGTAACTGCTTCACTGAGGCACTGTCTAC 524
DB 425 GGCCCTACCAAGAGTGGCTGTAACTGCTTCACTGAGGCACTGTCTAC 484
QY 525 TACTGTGAGATTGAGCATCCCCCACTGTGCAAGAGGTTGATCGCGCATGGT 584
DB 485 TACTGTGAGATTGAGCATCCCCCACTGTGAGAGAGCGCGCATGGT 544
QY 585 GTGAGCGAGTTGTAATTTGCACTCCCGAGCACTGAAATCTTGTGCTAACA 644
DB 545 GAGAGCGAGTTGTAATTTGCACTCCCGAGCGCTCTCTGAAATCTTGTGCTAACA 604
QY 645 TCTGTGTGGCTTCCCACTGAGGCACTGAGAGCACTGAGAGCACTGAGTGC 704
DB 605 TGAATGTGGCTTCCCACTGAGGCACTGAGAGCACTGAGAGCACTGAGTGC 664
QY 705 AGTTTGTGGCTTCCCACTGAGGCACTGAGAGCACTGAGAGCACTGAGTGC 764
DB 665 AGCTTGTGGCTTCCCACTGAGGCACTGAGAGCACTGAGAGCACTGAGTGC 724
QY 765 AACAGTCCCTACCCCGCGCATGCGCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 824
DB 725 GACAGCGCTTACCCCGCGCATGCGCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 784
QY 825 GTGCTGAGCGCTTACCCCGCGCATGCGCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 884
DB 785 GTGCTGAGCGCTTACCCCGCGCATGCGCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 844
QY 885 GTGCTGAGCGCTTACCCCGCGCATGCGCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 944
DB 845 GTGCTGAGCGCTTACCCCGCGCATGCGCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 904
QY 945 GGCACCTTCTACCTCTCTCACTGAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1004
DB 905 GGCACCTTCTACCTCTCTCACTGAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 964
QY 1005 ACGCTGATTAACCAATGAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1064
DB 965 ACGCTGATTAACCAATGAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1024
QY 1065 CCGAAGTGAAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1124
DB 1025 CCGAAGTGAAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1084
QY 1125 TACTATCAAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1184
DB 1085 TACTATCAAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1144
QY 1185 AACCGAAGTGAAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1244
DB 1145 AACCGAAGTGAAGTGGTCTGCGCATGAGTGGTCTGCGGGGAGAGCGCGA 1204
QY 1245 GGCCTCTGCAAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1304
DB 1205 GGCCTCTGCAAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1264
QY 1305 CAGTTTGTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1364
DB 1265 CAGTTTGTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1324
QY 1365 TACAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1424
DB 1325 TACAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1384
QY 1425 GAGAGTTCATGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1484
DB 1385 GAGAGTTCATGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1444
QY 1485 GCAAGTTCGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1544
DB 1445 GCAAGTTCGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1504
QY 1545 ACGTGAAGAAACAGTTCGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1604

DB 1505 ACGTGAAGAAACAGTTCGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1564
QY 1605 GCGGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1664
DB 1565 GCGGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1624
QY 1665 GCGGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1724
DB 1625 GCGGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1684
QY 1725 GAGAGGCTTCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1784
DB 1685 GAGAGGCTTCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1744
QY 1785 CAAAATGAGGCTTCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1844
DB 1745 CAAAATGAGGCTTCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1804
QY 1845 GATGAGTCCGATGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1904
DB 1805 GATGAGTCCGATGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1864
QY 1905 GTGCTGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1964
DB 1865 GTGCTGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1924
QY 1965 GTGCTGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2024
DB 1925 GTGCTGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 1984
QY 2025 GCTCATGCTTCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2084
DB 1985 GCTCATGCTTCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2044
QY 2085 TTCTGAGTCTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2144
DB 2045 TTCTGAGTCTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2104
QY 2145 AAAGTATCATCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2204
DB 2105 AAAGTATCATCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2164
QY 2205 CTGAGTGAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2264
DB 2225 CTGAGTGAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2284
QY 2285 GATGAGTTCATGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2344
DB 2345 GATGAGTTCATGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2304
QY 2385 ACGTGAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2444
DB 2345 ACGTGAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2404
QY 2445 AGTGAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2504
DB 2405 AGTGAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2464
QY 2505 GATGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2564
DB 2465 GATGAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2524
QY 2565 AAGCAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2624
DB 2525 AAGCAAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2584
QY 2625 GTATAGGCACTGAGGCACTGAGGCACTACCCCGCACTGAGTGGTCTGCGGGGAGAGCGCGA 2684

Db 2585 GTATAGGGGCGGG---GCCACCAAAATGTGTACACTGGGGGCGACCCATGTCACCC 2641
 QY 2685 CCTGATATCAGAGAGGAACTACTGACGACATTTATGCTGTGGCTTCCCCCCCCCAACACA 2744
 Db 2642 CCAGTGTGACGC-CCTGCGAGCTGTGAGACTGAGACCGCTACTGTGCACACAGGCC-CCAGA 2699
 QY 2745 ACCGACGCTGTGACTGCATCTCTTAGACTCAGAGT 2780
 Db 2700 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 2735

RESULT 7
 AAD13155
 ID AAD13155 standard; DNA; 3147 BP.
 XX AAD13155;
 AC
 XX 16-OCT-2001 (first entry)
 XX
 XX Human membrane-type serine protease (MTSP) 1 protease domain DNA.
 DE Human membrane-type serine protease (MTSP) 1 protease domain DNA.
 XX Human; transmembrane serine protease; membrane-type serine protease;
 KW MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;
 KW lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;
 KW matrilysin; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1665..2590
 CDS /tag= a
 FT /product= "Human transmembrane serine protease (MTSP) 1
 FT /note= "CDS does not include start codon"
 FT /partial
 PN WO200157194-A2.
 PD 09-AUG-2001.
 XX
 XX 02-FEB-2001; 2001WO-US003471.
 PF
 XX 03-FEB-2000; 2000US-0179982P.
 PR 18-FEB-2000; 2000US-0183542P.
 PR 22-JUN-2000; 2000US-0213124P.
 PR 26-JUL-2000; 2000US-0220970P.
 PR 08-SEP-2000; 2000US-0065798P.
 PR 22-SEP-2000; 2000US-0234840P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 XX Madison EL, Ong EO, Yeh J;
 PI
 XX WPI; 2001-488877/53.
 DR P-PSDB; AAB06936.
 PT Novel single chain polypeptide comprising protease domain of type-II
 PT membrane-type serine protease or its catalytically active portion useful
 PT for treating and preventing cancer and tumor.
 XX
 XX Example 6; Page 225-227; 256pp; English.
 PS
 XX The invention relates to transmembrane serine proteases and their
 CC corresponding nucleotides and the protease domain of a type-II membrane-
 CC type serine protease (MTSP). MTSP is useful for identifying compounds
 CC that modulate or inhibit its proteolytic activity and for formulating a
 CC medicament for treating neoplastic disease. MTSP and its corresponding
 CC nucleotides are useful in preventing or treating tumours or cancers such
 CC as lung carcinoma, colon adenocarcinoma and ovarian carcinoma, in
 CC diagnostics and in hybridisation assays. MTSP is useful as a diagnostic
 CC marker for tumour development, growth and/or progression and as
 CC immunogens to generate antibodies that specifically bind to it. MTSP DNA
 CC is useful in a yeast two-hybrid system and in gene therapy. The present

CC sequence is a DNA encoding protease domain of human MTSP1 protein (also
 CC called matrilysin)
 XX
 SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;
 Query Match 60.6%; Score 1883.2; DB 4; Length 3147;
 Best Local Similarity 81.2%; Pred. No. 0;
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

45 GATCGAGCCGCAAAACCATGGGTAGCAATCGGGGCGGACGCGGAGGGGCTCTCAG 104
 Db 5 GAGCGGCTCGGGGTACATGGGAGACGATCGGGCCGCAAGGCGGAGGGGCCGAAG 64
 QY 105 GACTTCGGCGCGGAGACTCAAGTACAATCCCGCTAGAGAACATGAAATGCTTTGAGAG 164
 Db 65 GACTTCGGCGCGGAGACTCAAGTACAATCCCGGACGAGAAAGTAAATGCTTGAGAGAA 124
 QY 165 GGTGTGAGATTCCTGCTCGTCAACAATGCGCAAGAAATGGAAGGAGGAGGCGCCAGGCG 224
 Db 125 GGTGTGAGATTCCTGCTCAAGTACAACAAGTCAAGAAAGTGAAGGAGATGACCCGAGGCG 184
 QY 225 TGGGTGTGTGTGGTGGAGTGTGCTGCTCACTTCTCTGCTCTCCCTCATGCTGTGCT 284
 Db 185 TGGGTGTGTGTGGAGCGCTGTGATCGGCTCTCTCTTGTGTGTGTGGGATCGGCTTC 244
 QY 285 CTGGTGTGCACTTCATTATCGAAATGTGGGCTTCAAAAGTCTTCAATGCGCATCTG 344
 Db 245 CTGGTGTGCAATTTGCAATACCGGAGCTGCTGTCAGAAAGTCTTCAATGCTCATG 304
 QY 345 AGGATCACAAAATGAAATCTTTCTGAAATGCTATGAGAACTCCACCTCCACAGATTATC 404
 Db 305 AGGATCACAAAATGAAATTTTGTGAATGCTTACGAAATCTCCAACTCCACTGATTTGTA 364
 QY 405 AGCCTGAGCGAGCGAGTGAAGAGGCGCTGAACTGTGATCAATGAATGCTGTCTGTG 464
 Db 365 AGCCTGAGCGAGAGTGAAGAGGCGCTGAACTGTGATCAAGCGGAGTCCCATTCCTG 424
 QY 465 GGTCCCTACCAAGAAAGTGGCTGTAACTGCTTCAGTGAAGGCAATGATCGCCTAC 524
 Db 425 GGCCTTACCAAGAGGAGTGGCTGTGAACGCTTCAGCAGAGGCGACGTCATCGCCTAC 484
 QY 525 TACTGTGCAAGTTTCAAGATCCGCCCAACACTGGGAGAAAGTTGATCGCGCATGCT 584
 Db 485 TACTGTGCAAGTTTCAAGATCCCGACAGACTGTGTGAGAGGCGGAGCGCTCATGCT 544
 QY 585 GTGAGAGCAATTTAAATTTGCAATGCAACCCCGAGCAAGGCACTGAATCTTGTGTGTA 644
 Db 545 GAGAGAGCGTAACTGATGATGCTGCCCCCGGCGGCGCTCCGAAAGTCTTGTGTGATC 604
 QY 645 TCTGTGTGAGCTTCCCATTTGACCCCAAGATGCTGCAAGAGACTCGAACAAGCTGC 704
 Db 605 TCAGTGTGCTTTCCCGACGAGCTCCAAAACAGTACAGAGAGCCGAGCAACAGCTGC 664
 QY 705 AGTTTTCCTCGATGCGCAATGCGGAGTGAAGAGTGAACAGCTTCACTACCCCTGCTTCCC 764
 Db 665 AGCTTTGCTCGACGCGCGCGGTGTGAAGTGAATGAGCTTCAACAAGCGCGCTTCCCT 724
 QY 765 AACAGTCCCTAACCGGAGCATGCGCGTGCAGTGGTCTGTGCGGAGGAGCGCGACTCT 824
 Db 725 GACAGCCCTTAACCCGCTCATGCGCGTGCAGTGGGCGCTGTGCGGAGGAGCGCGACTCA 784
 QY 825 GTGCTGAGCTTCACTTCCGAAAGCTTTGATGTGCTCTCTGTATGAGCAATGAGGAGTAC 884
 Db 785 GTGCTGAGCTTCACTTCCGAGCTTTGAACTTGTGCTGTGAGAGAGCGGAGCGAGC 844
 QY 885 CTGTGACCGGTATGATAGCTGAGCGCCATGGAACCCCAAGCTGTGTGTGTGTGT 944
 Db 845 CTGTGAGCGGTATGATAGCTGAGCGCCATGAGAGCCCAAGCTGTGTGTGTGTGTGT 904
 QY 945 GGCACCTTTCACCTCTCTTAACCTTGAACCTTCTCTCTCTCCAGAAAGTCTTCTTGTCTC 1004
 Db 905 GGCACCTTTCACCTCTCTTAACCTTGAACCTTCTCTCTCTCCAGAAAGTCTTCTTGTCTC 964

QY 1005 ACGGTGATTAACCACTACTGACCGGGGCAATCTGGCTTTGAGGCCACTTCTTCCAGCTG 1064
 DB 965 AACTGTGTTAACAACACTGAGCGGGGCAATCCGGCTTTGAGGCACTTCTTCCAGCTG 1024
 QY 1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTGTGATGACACCCAGGAGCAATTAGAGGCC 1124
 DB 1025 CTTAGAGATGAGCAGCTGTGGAGGCGCTTTAGTAAAGCCAGGGGACATTCAAGGCC 1084
 QY 1125 TACTATCCAGGCGCACTACCGGCCCAACATCACTGACATGGAATCAAGGTGCCAAC 1184
 DB 1085 TACTACCCAGGCGCACTACCCCAACATGACTGACATGGAACATGAGGTGCCAAC 1144
 QY 1185 AACCGGAAGTGAAGGTGGCTTCAAACTCTTATCTGTGTGACCCCAAGTCAAGT 1244
 DB 1145 AACCAAGATGTGAAGGTGAGCTTCAAAATCTTCTACTGTGTGAGCCCGGCTGGCC 1204
 QY 1245 GGTCTCTGACCAAGGACTATGTGAGATCAACGGGGAGAAATGACTGCGGTGAGGCTC 1304
 DB 1205 GGCACCTGCGCCCAAGGACTATGTGAGATCAATGGGGAGAAATATCTGCGAGAGAGTCC 1264
 QY 1305 CAGTTTGTGTGACGACACAGCAAGATTAACATTCACATTCCTCAATCTGATCACTG 1364
 DB 1265 CAGTTTGTGTGACGACACAGCAAGATTAACATTCACATTCCTCAATCTGATCACTG 1324
 QY 1365 TACACGAGACACCGGGTCTCTAGCTGATGCTCTCTCTAGACCTCCCAAGCCGCTGCCA 1424
 DB 1325 TACACGAGACACCGGGTCTCTAGCTGATGCTCTCTCTAGACCTCCCAAGCCGCTGCCA 1384
 QY 1425 GGGATTTGATGTGACAGACTGACGCTGATCCGAAGAACTGCGCTGCGACGCGCTG 1484
 DB 1385 GGGCAGTTTACGTCGCGGACGCGGGCGGTGATTCGGAAGAGAGCTGCGCTGTGATGCTG 1444
 QY 1485 GCAGACTGCGCGGATTTATGTATGATGAGCTTACTGCGATGCAATGCCACCCAGCTTC 1544
 DB 1445 GCGGACTGACCGGACCAAGCATGAGCTCAACTGCAATTTGCGACGCGCGCCAGCTTC 1504
 QY 1545 ACGTCAAAAACAGTGTCTGCAAGCCCTCTCTGCTGTGTGACAGTGTCAAGCTGT 1604
 DB 1505 ACGTCAAAAACAGTGTCTGCAAGCCCTCTCTGCTGTGTGACAGTGTCAAGCTGT 1564
 QY 1605 GGGGACGGAAGTGAAGAGAGGCTGACAGCTGTCCGCTGGAGATTTCAGGTTCACAT 1664
 DB 1565 GGAAGCAACAGCGAGAGAGGAGGTGCAATTTGCGGCCCAACCTTCAAGGTTCACAT 1624
 QY 1665 GGGAACTGTCTCTCTGACAGCCAGAAAGTGAATGGGAAGACACTGTGAGATGGCT 1724
 DB 1625 GGGAACTGTCTCTGAAAGCCAGAGATGCAATGGGAAGACACTGTGAGATGGCT 1684
 QY 1725 GACGAGGCTTCAATGACAGCGTGAATGCTCTGTGACCAAAATATCTTACCGCTGC 1784
 DB 1685 GACGAGGCTCTCTGCCCCAAGGTGAACGTGCTGCTTGAACCAACACACTTACCGCTGC 1744
 QY 1785 CAAAATGAGCTCTGTCTGAGCAAGGGCAACCTGAGTGTGATGGGAAGAGGACTGTAGC 1844
 DB 1745 CTCAATGAGGCTCTGTCTGAGCAAGGGCAACCTGAGTGTGAGAGGAGAGGACTGTAGC 1804
 QY 1845 GATGCTCCGATGAGAAAACTGTGACTGTGGCTGCGATCTTTTACCAACAGGCTGC 1904
 DB 1805 GACGCTCAGATGAGAAAGACTGCGACTGTGGCTGCGGCTCATTCACAGACAGGCTGCT 1864
 QY 1905 GTGGTGTGAGCAAGAGCGGACAGAGGCGAGAGTGGCCCTTGGCAGAGTGAAGCTCCAGCC 1964
 DB 1865 GTTGTGGGGGAGCGAATGCGGATGAGGCGAGTGGCCCTTGGCAGAGTGAAGCTTGAATGCT 1924
 QY 1965 CTGGCCCAAGGCGCACTTGTGTGGGCTCTGCTCATCTCTCTGACTGGCTGGTCTGCA 2024
 DB 1925 CTGGCCCAAGGCGCACTTGTGGGCTCTTCCCTCATCTCTCTCAACTGGCTGGTCTCTGCC 1984
 QY 2025 GCTCATTTGCTTTAGAGATGACAAAAATTTCAAGTCTCAGACTACAGATGAGAGCGCC 2084
 DB 1985 GCACACTGCTCATGATGAGACAGAGATTTCAAGTACTCAGACCCCAAGAGTGAAGCGCC 2044
 QY 2085 TTCTGGGTGTGTGAGCAAGAGCAAGGCAAGTGTCTGTGGGTGACAGAGCTGAAGCTC 2144

DB 2045 TTCTGGGTGTGTGACAGACAGAGAGCGCGCCCTGGGTGACAGAGCGCAAGCTC 2104
 QY 2145 AAACGATATCAATCCACACCTTCTCTCAATGATTTTCACTTGAATGACATGCGCTTG 2204
 DB 2105 AAGCGATATCTTCCACACCTTCTCTCAATGATTTTCACTTGAATGACATGCGCTTG 2164
 QY 2205 CTGAGCTGAGAAAGTGTGTGAGTACAGACCGGTGTGGGCCCACTTGGCTGGCTGAT 2264
 DB 2165 CTGAGCTGAGAAAGTGTGTGAGTACAGACCGGTGTGGGCCCACTTGGCTGGCTGAT 2224
 QY 2265 GCTACCATGTCTTCTCTCTGCGCAAGCCATCTGGGTGACAGAGCTGGGGGCAACAAAA 2324
 DB 2225 GCTTCCCATGTCTTCTCTCTGCGCAAGCCATCTGGGTGACAGAGCTGGGGGCAACAAAA 2284
 QY 2325 GAGGAGATACCGAGAGCGTGTATCTCTGACAGAGGATGATCTGTTGATCAACAGACC 2384
 DB 2285 TATGAGAGCACTGGCGCGCTGTATCTGCAAAAGGATGATCTGCGTCAATCAACAGACC 2344
 QY 2385 ACCTGTGAGAGACTCATGCGCGAGAGATCAACCCCAAGATGATGTGTGGGTTCTTC 2444
 DB 2345 ACCTGTGAGAGACTCATGCGCGAGAGATCAACCCCAAGATGATGTGTGGGTTCTTC 2404
 QY 2445 AATTGGAGTGTGACTCTCTGCGAGAGGTGAATCTGTGAGCCCTTGTCAAGCGCGAGAAA 2504
 DB 2405 AAGCGGCGGTGAGACTCTCTGCGAGAGGTGAATCTGTGAGCCCTTGTCAAGCGCGAGAAA 2464
 QY 2505 GATGGGCGAATTTCCAGGCTGTGTGTGTGACTGGGGGTGAAGGCTGCGCTGAGAGAAC 2564
 DB 2465 GATGGGCGAATTTCCAGGCTGTGTGTGTGACTGGGGGTGAAGGCTGCGCTGAGAGAAC 2524
 QY 2565 AAGCGAGCGTGTACAAAGGCTCCCTGTGATTCGGGACTGGATCAAAAGGCACTGGG 2624
 DB 2525 AAGCGAGCGTGTACAAAGGCTCCCTGTGATTCGGGACTGGATCAAAAGGCACTGGG 2584
 QY 2625 GTATGACAGATGACAGACAGCTGCAAAACCCCAAGGATGCCGACATGACACA 2684
 DB 2585 GTATGAGGCGCGGG---GCCACCAAAATGTGACACTGCGGGGCCAACCATGTGACAC 2641
 QY 2685 CCTGATACAGAGAGAGAACTGACGACATTTATGTGTGCTTCCCCCCCAAGACA 2744
 DB 2642 CCAATGTGACG-CCTGAGAGGTGAGACTGAGACCTGACATGACACAGCGCC-CAGAG 2699
 QY 2745 ACCGAGCTGTGAATGCAATCTTGAAGTCAAGT 2780
 DB 2700 ACATACCTGTGACTCAATCTCCAGGGCTCAAAAT 2735

RESULT 8
 AAD13113
 ID AAD13113 standard; DNA; 3147 BP.
 XX
 AC AAD13113;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human membrane-type serine protease (MTSP) 1 DNA.
 XX
 DE Human, transmembrane serine protease; membrane-type serine protease;
 KM MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;
 KM lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;
 KM matrix; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 23..2590
 FT /tag= a
 FT /product= "Human transmembrane serine protease (MTSP) 1"
 XX
 PD WO200157194-A2.
 XX 09-AUG-2001.


```

QY 1605 GGGAGCGAAGTGCAGAGAGGGCTGCAGCTGCTCTGGAGTTTCAAGTGTCCAA 1664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1565 GGAGACAGACGCAAGAGAGGGGTGCAGTTGTCCGCCACCAAGCTTCAGTGTCCAA 1624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1665 GGGAGTGTCTCCCTCAGAGCCAGAAAGTATATGAGAGAGCAACTGTGAGATGGGTCT 1724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1625 GGGAGTGTCTCCCTCAGAAAGCCAGAGCTGCATATGAGAGAGCAACTGTGAGAGCGGTCC 1684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1725 GAGAGAGCTTCATGTGCAGAGCTGAAATGTGTCTTTTGCACCAATATACCTACCGCTGC 1784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1685 GAGAGAGCTTCCTGCCCCAAGGTGAAAGCTGTGCTGCTGATTCACCAACACCTACCGCTGC 1744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1785 CAAAATGGGCTCTGTCTGAGCAAGGGCAACCTGAGTGTATGAGAGAGAGAGCTGTAGC 1844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1745 CTCAAATGGGCTCTGTCTGAGCAAGGGCAACCTGAGTGTATGAGAGAGAGAGCTGTAGC 1804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1845 GATGCTCCGATGAGAAAACTGTGACTGTGGCTCGATCCTTTTCCAAACAGGCTCCG 1904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1805 GAGGCTCAGATGAGAAAGAGCTGCAGCTGTGGCTGCGGTTCATTCAGAGAGAGGCTCGT 1864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1905 GTGTGTGGTGGACAGAAATGGGACAGAGGGGAGTGGCCCTGAGAGGTGAGCCCTCCAGCC 1964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1865 GTGTGTGGGAGGACAGATGCGAGTGAAGGGAGTGGCCCTGAGAGGTGAGCCCTCGATGCT 1924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1965 CTGGGCGCAGGGCCACTTGTGTGGGGCTGGCTCATCTCTCTGACTGGCTGTCTGTCA 2024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1925 CTGGGCGCAGGGCCACTGTGGGTGTCTCTCATCTCTCTCCAACTGGGTGTCTGTCC 1984
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2025 GCTCATTTGCTTTCAAGATGACAAAAATTTCAAGTACTCAGACTACAGATGTGAGAGCGCC 2084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1985 GCACACTGTCTACATGTATGACAGAGAGATTCAAGTACTCAGACCCACAGAGTGTGAGCGCC 2044
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2085 TTCTGTGGTCTGTGTGACCAAGAGCAAGCGAGTCCCTGTGGGGTGCAGAGCTGAGCTC 2144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2045 TTCTGTGGGCTGTGACAGAGCAAGCGAGCGAGCGCCCTGTGGGGTGCAGAGCTGAGCTC 2104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2145 AAACGATCATCAACCCAGCCCTTCTTCATGATTTTCACTTGTGACTATGACATTCGCTTGG 2204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2105 AAGCGATATCTCCACCCCTTCTTCATGACTTCACTTGTGACTATGACATTCGCTTGG 2164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2205 CTGAGACTGAGAGTGGTGTGAGTACAGACCGTGTGGCCCACTGCGTGCCTGAT 2264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2165 CTGAGACTGAGAGAACCGGCAAGTACAGTCTCAATGTGTGGCCCACTGCGTGCCTGAT 2224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2265 GCTACCCATGTCTTCTCTGTGCGCAAGGCGCATCTGGGTCAACAGCTGGGGGCAACAAA 2324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2225 GCTCCCATGTCTTCTCTGTGCGCAAGGCGCATCTGGGTCAACAGGCGCAACCCAG 2284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2325 GAGGAGAGTACCGGAGCGCTGATCTCTGCAAGAGGTTGAGATCCGTGTCTTCAACCAAGCC 2384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2285 TATGAGAGGACTGGCGCTGTATCTTGCACAAAAGGTTGAGATCCGTGTCTTCAACCAAGCC 2344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2385 AACTGTGAGGACCTCATGCTCGGAGAGATCAACCCAGCAATGATGTGTGGGTTTCTCTC 2444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2345 AACTGTGAGGACCTCTCTGCGGAGAGATCAACCCAGCAATGATGTGTGGGTTTCTCTC 2404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2445 AGTGGGGGTGTGACTCTCTGCGGAGGTTGATCTGTGTGCGCCCTTGTCAAGCGCGGAGAA 2504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2405 AGCGGCGGCTGTGACTCTCTGCGGAGGTTGATCTGTGTGCGCGCCCTTGTCAAGCGGAG 2464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2505 GATGGGCGCAATGTTCCAGGCTGTGTGTGAGTGTGGGTGAAAGCTGTGGCTTCAAGAGAAC 2564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2465 GATGGGCGCAATCTTCTCAAGCGGTGTGTGTGAGTGTGGGTGAAAGCTGTGGCTTCAAGAGAAC 2524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2565 AAGCCAGGCTGTGTACACAAAGGCTCCTGTAGTTGGGAGCTGTGATCAAAAGAGACACTGGG 2624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 AAGCCAGGCTGTGTACAAAGGCTCCTGTAGTTGGGAGCTGTGATCAAAAGAGACACTGGG 2584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2625 GTATATGCGATGTGACAGACAGCCGACCAAAACCCCAAGGAGTGGCCGACATGACA 2684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2585 GTATATGCGGCGCGG---GCAACCCAAATGTGTACACTGTGGGGGCAACCCATGTGCCACC 2641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2685 CCGTGTATGAGAGAGAGAACTGACGACATTTATGTGTGGCCCTCCCCCCCCCAACACA 2744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 2642 CCAGTGTGACAGC-CTGTGAGGCTGTGAGACTGTGAGACCCCTGACTGTGACAGGCCCC-CCAGA 2659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2745 ACCCAGACTGTGAGACTGCATCTCTTAGACTCTAGAGT 2780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2700 ACATACACTGTGAACTCAATCTCCAGGCTCCAAAT 2735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
AAL53444
ID AAL53444 standard; DNA; 3147 BP.
AC AAL53444;
XX
AC AAL53444;
XX
DT 12-DEC-2002 (first entry)
XX
DE Type II transmembrane serine protease 1 coding DNA SEQ ID No. 1.
XX
XX Cytoleatic; type-II membrane-type serine protease 7; MTS7; malignancy;
KW neoplastic disease; pre-malignant lesion; tumour; zymogen form; cancer;
KW malignant; enzyme; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200272786-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US007903.
XX
PR 13-MAR-2001; 2001US-0275592P.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Madison EL, Ong EO;
XX
DR WP1; 2002-732827/79.
XX
DR P-PSDB; AAO22929.
XX
PT New transmembrane serine protease 7 (MTSP7) polypeptide for diagnosing
PT neoplastic diseases, monitoring tumor progress or therapeutic
PT effectiveness, or identifying MTSP7 modulators for treating tumors or
PT cancers.
XX
PS Disclosure; Page 168-172; 184pp; English.
XX
CC The invention relates to a purified single or two-chain polypeptide,
CC which comprises the protease domain of a type-II membrane-type serine
CC protease 7 (MTSP7) or its catalytically active portion. The polypeptide
CC comprising MTSP7 is useful for detecting or diagnosing a neoplastic
CC disease, a pre-malignant lesion, a malignancy or other pathologic
CC condition in a subject. This polypeptide is also useful for monitoring
CC tumour (e.g. tumor of the breast, cervix, prostate, lung, ovary or
CC colon) progress and/or therapeutic effectiveness. The inhibitor of the
CC polypeptide or the inhibitor of the polypeptide's zymogen form is useful
CC for treating or preventing a neoplastic disease, or tumour initiation,
CC growth or progression, or a (pre-)malignant condition. The polypeptide or
CC polynucleotide is also useful for identifying modulators of MTSP7, which
CC may be used to treat cancers or tumours. This polynucleotide sequence
CC represents the coding DNA for a protein of the type-II membrane-type
CC serine protease 1 relating to the invention
XX
SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATCGACCGCCCAAAACCATGGGTAGCAATCGGGGCGCCAGAGCGGGGGCTCTAG 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 GAGGCGCTCGGGGATCATGAGGAGCATGTGGCGCCGCAAGGCGGAGGGGCCCGAAG 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 GACTTGGCGGGGACTCAAGTACAACTCCCGGCTTAGAGAAATGATGCTTTGAGAG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


QY 2325 GAGGAGGTACCGGAGCGGTGATCTTCCAGAAAGGGTGAATCCGTGTCAATCAACGAGCC 2384
 DB 2285 TATGGAGGACATGCGCGGTGATCTTCCAGAAAGGGTGAATCCGTGTCAATCAACGAGCC 2344
 QY 2385 ACCGTGAGGACCTTCATGCGGAGCAGATCAACCCGAAATGATGTTGGTTCCTC 2444
 DB 2345 ACCTGGAGAACTCTCTGCGGAGCAGATCAACCCGAAATGATGTTGGTTCCTC 2404
 QY 2445 AGTGGGGGTGTGATCTTCCAGAGGTGAATCTGTGAGGCGCTTGTCAAGCGGAGAAA 2504
 DB 2405 AGCGGCGGTGATCTTCCAGAGGTGAATCTGTGAGGCGCTTGTCAAGCGGAGAAA 2464
 QY 2505 GATGGGCGAATGTTCCAGAGGTGTGTGTGATGCTGGGGTGAAGGCTGCGTCAAGAGAA 2564
 DB 2465 GATGGGCGAATCTTCCAGAGGTGTGTGTGATGCTGGGGTGAAGGCTGCGTCAAGAGAA 2524
 QY 2565 AAGCCAGGCGGTGATCAACAGGCTCCGTGATGTCGGGACCTGATCAAAAGACACTGGG 2624
 DB 2525 AAGCCAGGCGGTGATCAACAGGCTCCGTGATGTCGGGACCTGATCAAAAGACACTGGG 2584
 QY 2625 GTATAGACCATGAGCAGACAGCCGACCAACCAACCAAGGATGCCGACATGACCA 2684
 DB 2585 GTATAGGCGCGCGG---GCCACCCAAATGTATACATGCGGGGCCACCATGTGTACC 2641
 QY 2685 CCTGATACAGAGAGAGAACTGAGACATTTATGCTGTGCGCTTCCCCCAACACA 2744
 DB 2642 CCAATGTGTACG-CCTGACGCGCTGAGAGCTGAGACCGCTGATGCAACAGCGCC-CAGA 2699
 QY 2745 ACCCAGACTGTGAATGATCTCTTAGAAGTCAAGAGT 2780
 DB 2700 ACATACACTGTGAATCAATCTCCAGGCTCCAAAT 2735

RESULT 10
 AAL53445
 ID AAL53445 strand; DNA; 3147 BP.
 XX
 AC AAL53445;
 XX
 DT 12-DEC-2002 (first entry)
 XX
 DE Type II transmembrane serine protease 1 domain DNA SEQ ID No 3.
 XX
 KM Cytostatic; type-II membrane-type serine protease 7; MTSPT; malignancy;
 KM neoplastic disease; pre-malignant lesion; tumour; zymogen form; cancer;
 KM malignant; enzyme; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200272786-A2.
 PD 19-SEP-2002.
 XX
 PF 13-MAR-2002; 2002WO-US007903.
 XX
 PR 13-MAR-2001; 2001US-0275592P.
 XX
 PA (CORV-) CORVAS INT INC.
 P1 Madison EL, Ong EO;
 XX
 DR WPI: 2002-732827/79.
 DR P-PSDB; AAO22930.
 XX
 PT New transmembrane serine protease 7 (MTSP7) polypeptide for diagnosing
 PT neoplastic diseases, monitoring tumor progress or therapeutic
 PT effectiveness, or identifying MTSPT modulators for treating tumors or
 PT cancers.
 XX
 PS Disclosure; Page 174-176; 184pp; English.
 XX
 CC The invention relates to a purified single or two-chain polypeptide,

CC which comprises the protease domain of a type-II membrane-type serine
 CC protease 7 (MTSP7) or its catalytically active portion. The polypeptide
 CC comprising MTSPT is useful for detecting or diagnosing a neoplastic
 CC disease, a pre-malignant lesion, a malignancy or other pathologic
 CC condition in a subject. This polypeptide is also useful for monitoring
 CC tumour (e.g. tumor of the breast, cervix, prostate, lung, ovary or
 CC colon) progress and/or therapeutic effectiveness. The inhibitor of the
 CC polypeptide or the inhibitor of the polypeptide's zymogen form is useful
 CC for treating or preventing a neoplastic disease, or tumour initiation,
 CC growth or progression, or a (pre-)malignant condition. The polypeptide or
 CC polynucleotide is also useful for identifying modulators of MTSPT, which
 CC may be used to treat cancers or tumours. This polynucleotide sequence
 CC represents the coding DNA for a protein domain of the type-II membrane-
 CC type serine protease 1 relating to the invention
 XX
 SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
 Best Local Similarity 81.2%; Pred. No. 0;
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATGGGACCGGCAAAACCAATGGGTGACATGCGGGGCGGCAAGGCGGAGGCTTCAG 104
 DB 5 GAGCGGCTCGGGGTACCATGCGGAGCATGCGGCGGCAAGGCGGAGGCGGCGGAG 64
 QY 105 GACTTCGGCGCGGAGCTCAAGTACAACTCCGGCTAGAGACATGAATGCTTTGAGAG 164
 DB 65 GACTTCGGCGGCGGAGCTCAAGTACAACTCCGGCGGCAAGGAGTGAATGCTTTGAGAG 124
 QY 165 GGTGTGAGATCTCTGCTCGGCAAAATGCGGCAAGGAGGAGGCGGCGGAGCGCC 224
 DB 125 GCGGTGAGATCTCTGCTCGGCAAAATGCGGCAAGGAGGAGGAGGAGGCGGCGGCG 184
 QY 225 TGGGTGAGTGTGCGGTGCGGTGCTTCAAGCTTCTTCTCTCTCTCTCTCTCTCTCT 284
 DB 185 TGGGTGAGTGTGCGGTGCGGTGCTTCAAGCTTCTTCTCTCTCTCTCTCTCTCTCTCT 244
 QY 285 CTGCTGTGCGCACTTCATTTATGGAATGCGGCGGTTCAAAAGTCTTCAATGCGCATC 344
 DB 245 CTGCTGTGCGCACTTCATTTATGGAATGCGGCGGTTCAAAAGTCTTCAATGCGCATC 304
 QY 345 AGATCACAAAATGAGATCTTTCTGAGTGCATGAGAACTCCACCTCCACAGAGTTATC 404
 DB 305 AGATCACAAAATGAGATCTTTCTGAGTGCATGAGAACTCCACCTCCACAGAGTTATC 364
 QY 405 AGCCTGCGCACCGAGTGAAGAGCGGCTGAAAGCTGCTGTACATGAAGTCTCTCTCT 464
 DB 365 AGCCTGCGCACCGAGTGAAGAGCGGCTGAAAGCTGCTGTACAGAGGAGTCCCATCTCT 424
 QY 465 GATCCTTACCAAGAAAGTGGCTTGAATGCTTCACTGAGGAGGAGTCAATGCGCTTAC 524
 DB 425 GATCCTTACCAAGAAAGTGGCTTGAATGCTTCACTGAGGAGGAGTCAATGCGCTTAC 484
 QY 525 TACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 584
 DB 485 TACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 544
 QY 585 GTGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 644
 DB 545 GTGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 604
 QY 645 TCTGTGTGCGCTTCCCATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 704
 DB 605 TCTGTGTGCGCTTCCCATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 664
 QY 705 AGTTTGTGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 764
 DB 665 AGTTTGTGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 724
 QY 765 AAGAGTCCCTTACCGGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 824
 DB 725 AAGAGTCCCTTACCGGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 784

QY 825 GTGCTGAGCGCTGACCTTTCGGAAGCTTTGAGTGTGCTCCCTGTGATGAGATGAGCACTGAGTAC 884
 |||||
 Db 785 GTGCTGAGCGCTGACCTTTCGGAAGCTTTGAGTGTGCTCCCTGTGATGAGATGAGCACTGAGTAC 844
 |||||
 QY 885 CTGCTGACCGGTGTATGATGATGAGCTGAGCGCCCATGGAACCCCAAGCTGTGTGCTGTGT 944
 |||||
 Db 845 CTGCTGACCGGTGTATGATGATGAGCTGAGCGCCCATGGAACCCCAAGCTGTGTGCTGTGT 904
 |||||
 QY 945 GGCACCTTCTACCCCTCTTACCACTGACCTTCTCTCCCTCCCAAGAGCTTCTCTTTC 1004
 |||||
 Db 905 GGCACCTTCTACCCCTCTTACCACTGACCTTCTCTCCCTCCCAAGAGCTTCTCTTTC 964
 |||||
 QY 1005 AGCTGATTAACCAATATCTGACCGGCGACATCTGCGCTTGTGAGCGCACTTCTTCCAGCTG 1064
 |||||
 Db 965 AGCTGATTAACCAATATCTGACCGGCGACATCTGCGCTTGTGAGCGCACTTCTTCCAGCTG 1024
 |||||
 QY 1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTGTGATGACACCCAGAGGACATTTAGCAGCCCC 1124
 |||||
 Db 1025 CTTAGAGATGAGCAGCTGTGGAGGCGCTTACGTAAAGCCAGGGGACATTCACAGCCCC 1084
 |||||
 QY 1125 TACTATCCAGGCGCATACCCGCCCAACATCACTGACATGAGAAATTCAGAGTCCCAAC 1184
 |||||
 Db 1085 TACTATCCAGGCGCATACCCGCCCAACATCACTGACATGAGAAATTCAGAGTCCCAAC 1144
 |||||
 QY 1185 AACCGGAACGTGAAGGTGCGCTTCAAACTCTTATCTGTGTGAGACCCCAAGTACAGAGT 1244
 |||||
 Db 1145 AACCGGAACGTGAAGGTGCGCTTCAAACTCTTATCTGTGTGAGACCCCAAGTACAGAGT 1204
 |||||
 QY 1245 GAGCTCTGACACCAAGGATATGTGTGAGATCAAGGGGAGAAATCTGCGGTGAGAGTCC 1304
 |||||
 Db 1205 GAGCTCTGACACCAAGGATATGTGTGAGATCAAGGGGAGAAATCTGCGGTGAGAGTCC 1264
 |||||
 QY 1305 CAGTTTGTGTGAGCAGCAACAGCAGAAATTAACATCACTTCCATTTGATCACTG 1364
 |||||
 Db 1265 CAGTTTGTGTGAGCAGCAACAGCAGAAATTAACATCACTTCCATTTGATCACTG 1324
 |||||
 QY 1365 TACACGGAACACCGGCTCTAGCTGAGTACCTCTCTACGACTCCAAAGCAAGCCGTGCCA 1424
 |||||
 Db 1325 TACACGGAACACCGGCTCTAGCTGAGTACCTCTCTACGACTCCAAAGCAAGCCGTGCCA 1384
 |||||
 QY 1425 GGGATGTTCAATGTGCAAGACTGACGCTGCATCCGAAAGAACTGCGCTGCGAGCGCTGG 1484
 |||||
 Db 1385 GGGATGTTCAATGTGCAAGACTGACGCTGCATCCGAAAGAACTGCGCTGCGAGCGCTGG 1444
 |||||
 QY 1485 GCGACACTGCCGATTTATGATGAGAGCTTACGCGGATGCAATGACCAACAGCTTC 1544
 |||||
 Db 1445 GCGACACTGCCGATTTATGATGAGAGCTTACGCGGATGCAATGACCAACAGCTTC 1504
 |||||
 QY 1545 ACGTGCAAAAAACAGTTCTGTCAAGCCCTCTTCTGTGTGTGACAGTGTCAAGCTGT 1604
 |||||
 Db 1505 ACGTGCAAAAAACAGTTCTGTCAAGCCCTCTTCTGTGTGTGACAGTGTGTCAAGCTGT 1564
 |||||
 QY 1605 GGGGACGGAAGTGAAGAGAGGAGCTGACGCTGTCTGTGAGAGTTTCAAGTGTTCAT 1664
 |||||
 Db 1565 GGGGACGGAAGTGAAGAGAGGAGCTGACGCTGTCTGTGAGAGTTTCAAGTGTTCAT 1624
 |||||
 QY 1665 GGGGAGTGTCTCTCCAGAGCAGAAAGTGTATGGGAAAGCACTGTGAGAGTGGCT 1724
 |||||
 Db 1625 GGGGAGTGTCTCTCCAGAGCAGAAAGTGTATGGGAAAGCACTGTGAGAGTGGCT 1684
 |||||
 QY 1725 GACGAGGCTTATGTGACAGCGTGAATGTCTCTTGTGACCAAAATATACCTACCGCTGC 1784
 |||||
 Db 1685 GACGAGGCTTATGTGACAGCGTGAATGTCTCTTGTGACCAAAATATACCTACCGCTGC 1744
 |||||
 QY 1785 CAAATATGCTCTGTCTGTGAGAGAGGCAACCTGTGATGTGATGGGAAAGCAAGCTGTAGC 1844
 |||||
 Db 1745 CTAATATGCTCTGTCTGTGAGAGAGGCAACCTGTGATGTGATGGGAAAGCAAGCTGTAGC 1804
 |||||
 QY 1845 GATGAGCTCCGATGAGAAATCTGTGACCTGTGGGCTGCGATCCTTTACCAAAACAGGCTGC 1904
 |||||
 Db 1805 GATGAGCTCCGATGAGAAATCTGTGACCTGTGGGCTGCGATCCTTTACCAAAACAGGCTGC 1864
 |||||
 QY 1905 GTGCTGTGTGACGAATGTGAGAGGAGCGAGTGTGCTGTGAGAGCTTCCAGCC 1964
 |||||

Db 1865 GTTGTGTGGGGACAGGATGCGATGAGGGCGAGTGGCCCTGTGAGGATGAGCTTGCATGCT 1924
 |||||
 QY 1965 CTGGGCGAAGGCGCATCTGTGTGGGGCTGTGATCATCTCTTGACTGAGCTGTCTGTGA 2024
 |||||
 Db 1925 CTGGGCGAAGGCGCATCTGTGTGGGGCTGTGATCATCTCTTGACTGAGCTGTCTGTGA 1984
 |||||
 QY 2025 GCTCATTTGCTTTCAGAGTGAACAAAATTTCAAGTACTCAGACTCAACGATGTGAGCGCC 2084
 |||||
 Db 1985 GCACACTGTATCATGTATGACAGAGATTCAGGTACTCAGAACCCCAAGCAGTGTGAGCGCC 2044
 |||||
 QY 2085 TTCTGTGGTGTGTGTGACCAAGAGCAAGCGGAGCTCTGTGGGTGTGAGAGCTGAAAGCTC 2144
 |||||
 Db 2045 TTCTGTGGTGTGTGTGACCAAGAGCAAGCGGAGCTCTGTGGGTGTGAGAGCTGAAAGCTC 2104
 |||||
 QY 2145 AAAGTATCATCAACCAACCTTCTCTCAATGATTTTCACTTGCATATGACATGTGCTTGT 2204
 |||||
 Db 2105 AAAGTATCATCAACCAACCTTCTCTCAATGATTTTCACTTGCATATGACATGTGCTTGT 2164
 |||||
 QY 2205 CTGAGACTGAGAAAGTGTGTGAGTACAGACCGTGTGTGCGCCCATCTGTCTGTGAT 2264
 |||||
 Db 2165 CTGAGACTGAGAAAGTGTGTGAGTACAGACCGTGTGTGCGCCCATCTGTCTGTGAT 2224
 |||||
 QY 2265 GCTTACCATGTCTTCTCTGTGTGAGAGCGCATGTGTGTGACAGGCTGTGGGGGCAACAAAA 2324
 |||||
 Db 2225 GCTTACCATGTCTTCTCTGTGTGAGAGCGCATGTGTGTGACAGGCTGTGGGGGCAACAAAA 2284
 |||||
 QY 2325 GAGGAGGTACCGGAGCGCTGATCTGTGAGAAAGGTGAGATCCGTGTCAACCAAGACC 2384
 |||||
 Db 2285 TATGTGAGGCACTGTGCGCGCTGTGATCTGTGAAABGGTGTATCCGCGTCAACCAAGACC 2344
 |||||
 QY 2385 ACCTGTGAGGACCTCATGTGCGAGAGATCAACCCCAAGAAATGATGTGTGTGCTTCTC 2444
 |||||
 Db 2345 ACCTGTGAGGACCTCATGTGCGAGAGATCAACCCCAAGAAATGATGTGTGTGCTTCTC 2404
 |||||
 QY 2445 AGTGTGGGTGTGATCTGTGCGAGGATCACTGTGTGTGCGCCCTTGTCAAGCGCGAGAAA 2504
 |||||
 Db 2405 AGTGTGGGTGTGATCTGTGCGAGGATCACTGTGTGTGCGCCCTTGTCAAGCGCGAGAAA 2464
 |||||
 QY 2505 GATGGGCGAATGTTCACGAGCTGTGTGTGTGAGTGTGAGGCTGTGAGAGAAC 2564
 |||||
 Db 2465 GATGGGCGAATGTTCACGAGCTGTGTGTGTGAGTGTGAGGCTGTGAGAGAAC 2524
 |||||
 QY 2565 AAGCAGGCGGTGTACACAAAGCTCTCTGTGTGTGTGAGTGTGAGTCAAAAGGACACTGGG 2624
 |||||
 Db 2525 AAGCAGGCGGTGTACACAAAGCTCTCTGTGTGTGTGAGTGTGAGTCAAAAGGACACTGGG 2584
 |||||
 QY 2625 GTATAGCAGATGAGACAGACCGCAACCAACCAAGGATGCGCGACATGACA 2684
 |||||
 Db 2585 GTATAGGAGGCGCGG--GCCAACCAGATGTGTACACTGTGCGGCGCACCTCATGTCAAC 2641
 |||||
 QY 2685 CTTGATATCAGAGAGAGAAACATGACGACATTTATGTGTGTGCTTCCCTCCCAAGACAA 2744
 |||||
 Db 2642 CCAATGTGCAAG-CCTGCAAGCTGTGAGACTGACCGCTGACTGTGACCAAGCGCC-CGAGA 2699
 |||||
 QY 2745 ACCCAAGCTGTGAATGTGATCTTTAGACTGTAGAGT 2780
 |||||
 Db 2700 ACATACACTGTGAATCTCAAGGAGCTTCCAAAT 2735
 |||||

RESULT 11

ABZ22450
 ID ABZ22450 standard; cDNA; 3147 BP.

XX ABZ22450;

XX AC 24-MAR-2003 (first entry)

DE Human membrane-type serine protease MTSPI encoding cDNA SEQ ID NO:1.

XX Human: membrane-type serine protease; enzyme; MTSPI0; cytosolic;

KW type-II membrane-type serine protease; neoplastic disease; tumour; MTSPI;

KW matrilipase; gene; ss.

QY	1365	TACACGGACACACCGGGTCTCAGCTGAGTAACTCTCTCAGACATCCAAACGACCCGTGGCCA	1424
Db	1325	TACCCGACACACCGGCTCTTAGCTGAAATACCTCTCTCAACATCTCCATGACCTCATGCCG	1384
QY	1425	GGGATGTTTCATATGTCACAGACTGGAACGGATGATCCGAAAGGAATCGCGCTGCGA	1484
Db	1385	GGGAGATTCAAGTGCACCGAAGGGGCGGTGTATCCGGAAGGAGCTGGCTGTGATGAGCTGG	1444
QY	1485	GCAGAATGCCCCGGATTTATGTATGATGAGGTTTACTGCGGATGCAATGGCACCCACCAATTC	1544
Db	1445	GCCCACTGCAACCGACCCACAGCGATGAGTCAACTGCAAGTTTGCAGACGCCGCGCACCAATTC	1504
QY	1545	AACGGCAAAAACGATTTCTGCAACCCCTCTTGGGCTGTGACAGTGTCAACAGACTGT	1604
Db	1505	ACGTGCAAGAACAAAGTTTGTGAACCCCTCTTCTGGGCTGTGCAAGTGTGAACAGACTGC	1564
QY	1605	GGGAGCGGAATGACGAGAGGGGCTGACAGCTGTCTCTGCGGAGTTTCAAGTTTCCAAAT	1664
Db	1565	GGAGACAAACACCGACGAGACAGGGGTGCAAGTTGTCTCGGCCAGACCTTCAAGTGTTCAAAT	1624
QY	1665	GGGAAATGTCTTCCCTCAGAGCCAGAGTTTATGTGGAAGCAACTGTGGAAGATGGGTCT	1724
Db	1625	GGGAAATGCTCTCTCGAAGAACACAGTGCATATGGGAAGACAGACTGTGGGGAAGGGGTTC	1684
QY	1725	GACGAGGTTTCATGTGACAGGGTGAATGTGCTCTTTGCAACAAATPATCTACCGCTGC	1784
Db	1685	GACGAGGCTCTCTGCCCCCAAGGTGAACGTGTCTCTTTACCAACACACTTACCGCTGC	1744
QY	1785	CAAAATGCGCTCTGTCTGAGCAAGGGACACCCCTGACATGTGTATGTGGGAACGCACTGTAGC	1844
Db	1745	CTCATATGGGCTCTGTGTAGGACAGGGGCAACCTGTAGTGTGACGGGAAGAGAGACTGTAGC	1804
QY	1845	GATGAGCTCCGATGAGAAAACTGTGACTGTGAGCTGTGCATCTTTTACCAACAGGCTGCG	1904
Db	1805	GACGCGCTCAGATGAGAAAGCACTGCGCACTGTGGGCTGCGGTCAATTCACACAGCAAGCTGTGT	1864
QY	1905	GTTGTTGTGTGTGACGAATCGGACGAGAGGGGCGAGTGGCCCTGCGACAGTGAAGCTGTCAAGCC	1964
Db	1865	GTTGTTGGGGGACGAGATGCGGATGAGGACAGATGGCCCTGTGACAGATGAGGCTGTGATGTCT	1924
QY	1965	CTGGGCGCAAGGCGCACTTGTGTGTGAGGCGCTGTGCTCATCTCTCTGACTGTGCTGTGTCA	2024
Db	1925	CTGGGCGCAAGGCGCACTCTGGGCGTCTTCTCTCATCTCTCCCACTGGCTGTGTCTGTGCC	1984
QY	2025	GCTCATTTGCTTTCAGAGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGACGCGC	2084
Db	1985	GCACACTGTGTCATGATGACAGAGATTCAGGTACTCAGACCCACGACAGTGAACGCGCC	2044
QY	2085	TTCTCTGGGCTCTCTGTGACCAAGAGCAAGGCGAGTGTCTGGGGTGTGCAAGACTGAAGTCTC	2144
Db	2045	TTCTCTGGGCTTTCACAGACCAAGGACGAGCGACGCGCCCTCTGGGGTGTGCAAGAGCAAGGCTTC	2104
QY	2145	AAACGTATCATCAACCAACCTTCTCTTCATGTATTTCACTTTCGACTATGATGACGTGCTTG	2204
Db	2105	AAGCGATATATCTCCACCCCTTCTTCAATGACTTCACTTTCGACTATGATGACGTGCTGTG	2164
QY	2205	CTGGAGCTTGGAGAAAGTGTGGTGAATGACAGACGTGTGTGCGCCCATCTGCTGTGCTGAT	2264
Db	2165	CTGGAGCTTGGAGAAACCGGCAAGTACAGTCAATGATGTGCGGCCCATCTGTCTGTGCGGAC	2224
QY	2265	GCTACCCATGTCTTCCCTGTCTGTGCAAGGCGCATCTGGGTTCAGAGGTGGGGGCGACCAAAA	2324
Db	2225	GCCTCCCATGTCTTCCCTGTCTGTGCAAGGCGCATCTGGGTTCAGAGGTGGGGGCGACCCAG	2284
QY	2325	GAGGAGGATACCGGAGCGCTGATCTGTGCAAGAGGGTGAGATCCGTGTATCAACAGACCC	2384
Db	2285	TATGAGAGCACTGGCGCGCTGTGATCTGTGCAAAAAGGGTGAGATCCGCTATCAACAGACCC	2344
QY	2385	ACCTGTGAGAACCTCAATGCGGACGACGATCAACCCACGAATGATGTGTGGGTTTCTC	2444
Db	2345	ACCTGTGAGAACCTCTGCGCGAGGACGATCAACCGCGCATGATGTGCGTGGGCTTCTCTC	2404
QY	2445	AGTGGGGGTGTGAGCTCTGTCCACAGGGTGACTGTGTGTGCGCCCTTGTCAAGCGCGAGAAAA	2504

Db	2405	AGCGGCGCGTGAGACTCCTGCACAAGGATGATTCGCGGGGAGACCCCTGTCCAGGTGAAAGCG	2464
QY	2505	GATGGGGGAATGTTCCAGGCTGTGTGTAGCGTGGGGTGAAGGCTGGGCTCAGAGAAC	2564
Db	2465	GATGGGGGGAATCTTCCAGGCCGGGTGTGGTAGCTGGGGAGACGGCTGGCTCAGAGAAC	2524
QY	2565	AAGCCAGGCGGTGACACAAGGCTCCCTGTAGTTGGGACTGGATCAAGAGACACTGGG	2624
Db	2525	AAGCCAGGCGGTGACACAAGGCTCCCTGTGTGGGACTGGATCAAGAGACACTGGG	2584
QY	2625	GTATAGCAGCATGGACAGACGCCGACCAACAACCACAGGGATGCCCGATGACACA	2684
Db	2585	GTATAGGGGCGCGG---GCCACCCAAATGTGTACACTCTGCGGGGCAACCCATGTGCCAC	2641
QY	2685	CCTGGATPCAGAGAGAGAACCTGACGACATTTATGTGCTGTGGCCCTCCCCCCCCCAACACA	2744
Db	2642	CCAGGTGTGCAAG-CCTGCAGGCTGGAGCTGGACCGCTGACTGCACAGCGCC-CCAGA	2699
QY	2745	ACCCAGACTGAACTGCATCCTTAGACTCAGAGT	2780
Db	2700	ACATPCACTGTGAATCAATCTCCAGGGGCTCCAAT	2735

RESULT 12
ABZ22451
ID ABZ22451 standard; cDNA; 3147 BP

AC ABZ22451;

DT 24-MAR-2003 (first entry)

Human MTSPI protease domain encoding cDNA SEQ ID NO:3.

KW Human; membrane-type serine protease; enzyme; MTS10; cytosolic;
KW type-II membrane-type serine protease; neoplastic disease; tumour; MTS1;
KW matrix; gene; ss.

OS Homo sapiens

FH	Key	Location/Qualifiers
FT	CDS	1865. .2590

```

/ product= "MTSP1 protease domain"
/ note= "no start codon given"

```

PN WO200292841-A2.

PD 21-NOV-2002

PF 14-MAY-2002; 2002WO-US015332.

PR 14-MAY-2001; 2001US-0291001P.

PA (CORV-) CORVAS INT INC.

PI Madison EL, Yeh J;

DR WPI; 2003-129309/12.

XX

PT type serine protease

PT treating such diseases.

PS Example 1; Page 183-185; 198pp; English

The present invention describes a polypeptide comprising a purified single or two chain polypeptide, which comprises the protease domain of type-II membrane-type serine protease (MTSP10) or its catalytically active portion, or a mutant of it, where up to 50 % of the amino acids

CC are replaced with another amino acid, and the resulting polypeptide is a
 CC single chain or two chain polypeptide that has a catalytic activity of at
 CC least 1-10 % of the unmutated polypeptide. MTSPI0 has cytostatic
 CC activity. The polypeptide containing the protease domain of the MTSPI0 is
 CC useful for detecting a neoplastic disease, and for diagnosing the
 CC presence of a pre-malignant lesion, a malignancy, or other pathologic
 CC condition in a subject or monitoring tumour (e.g. breast, cervix,
 CC prostate, lung, ovary or colon tumour) progression and/or therapeutic
 CC effectiveness. An inhibitor of the polypeptide containing the protease
 CC domain of MTSPI0 is useful for treating or preventing neoplastic disease
 CC in a mammal. An inhibitor of the activation cleavage of the zymogen form
 CC of the MTSPI0 polypeptide is useful for inhibiting tumour initiation,
 CC growth or progression, or treating (pre-)malignant conditions of the e.g.
 CC breast, cervix, prostate, lung, ovary or colon. The present sequence
 CC encodes the protease domain of human MTSPI (also known as matrilase),
 CC which is used in an example from the present invention

XX Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Query Match 60.6%; Score 1883.2; DB 7; Length 3147;
 Best Local Similarity 81.2%; Pred. No. 0;
 Matches 223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATCGAGCCGCCAAACCATGGTAGCAATCGGGGCGGCAAGGCCGGAAGGGGCTCTCAG 104
 Db 5 GAGCGGCTTCGGGGTACCATGGGAGGATCGGGCCCGCAGGGCGAGGGGCCCGGAAG 64
 QY 105 GACTTGGGCGGGGACTCAAGTACAACTCCGCTAGAGAACATGATGCTTTGAGAG 164
 Db 65 GACTTGGGCGGGGACTCAAGTACAACTCCGCGCAGAGAAAGTATGGCTTGGAGGAA 124
 QY 165 GGTGTGGAATTTCTGCTCTCGGAAACAATGCCAAGAAAGTGAAGAGGAGGCCCGCAGGCG 224
 Db 125 GGCCTGGAATTTCTGCTCTCGGAAACAACGTCAGAAAGTGAAGAGTGGCCCGGCGCG 184
 QY 225 TGGGATGCTGCTGATGGCAGTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 284
 Db 185 TGGGATGCTGCTGATGGCAGTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 244
 QY 285 CTGGTGTGGACCTTCATTAATGGAATGCGGGTTCAGAAAGTCTTCAATGAGCCATCTG 344
 Db 245 CTGGTGTGGACCTTCATTAATGGAATGCGGGTTCAGAAAGTCTTCAATGAGCCATCTG 304
 QY 345 AGGATCACAAAATGATGATCTTTCTGATGCGTATGAAACTCCACTCCACAGATTTATC 404
 Db 305 AGGATCACAAAATGATGATTTTGTGATGCTTACGAAACTCCACTCCAGTTTGTGA 364
 QY 405 AGCTGGCCAGCAGATGGAAGAGGCGTGAAGCTGCTGTATGAATGAATGCTCTGTCTG 464
 Db 365 AGCTGGCCAGCAGATGGAAGAGGCGTGAAGCTGCTGTATGAAGCTGCTGTCTG 424
 QY 465 GGTCCCTTACCAAGAAAGTGGCTGTAACTGCTTCAGTGAAGGAGAGTGTCACTGCTAC 524
 Db 425 GGCCTCTACCAAGAAAGTGGCTGTGAAGCTTCAGAGAGGAGAGTGTCACTGCTAC 484
 QY 525 TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 584
 Db 485 TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
 QY 585 GTGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
 Db 545 GAGAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604
 QY 645 TCTGTGTGGCTTCCCATTTGACCCGAGAGTGTGAGAGAGTCAAGAGCAACAGTGC 704
 Db 605 TCAATGTGTGGCTTCCCATTTGACCCGAGAGTGTGAGAGAGTCAAGAGCAACAGTGC 664
 QY 705 AGTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764
 Db 665 AGTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
 QY 765 AACAGTCCCTACCGGCGCATGCGCTGCGAGTGGATCTGCGGGGGAGAGCGCGAGCTCT 824

Db 725 GACAGCCCTTACCCCGCTCATGCCGCTGCAGAGGGCCCTGCGGGGAGAGCGCGACTCA 784
 QY 825 GTGCTGAGCTCACTTCCGAACTTTGATGTGCTTCTGTATGATGATGATGATGATGATGAT 884
 Db 785 GTGCTGAGCTCACTTCCGAACTTTGATGTGCTTCTGTATGATGATGATGATGATGATGAT 844
 QY 885 CTGGTACCGTGTATATAGCTGAGGCCCATGAGAACCCGAGCGTGTGTGTGTGTGTGTGT 944
 Db 845 CTGGTACCGTGTATATAGCTGAGGCCCATGAGAACCCGAGCGTGTGTGTGTGTGTGTGT 904
 QY 945 GGCACCTTCTGACCTTCTTCAAACTGACCTTCTCTCTCTCCGAGAACCTTCTCTGTG 1004
 Db 905 GGCACCTTCTGACCTTCTTCAAACTGACCTTCTCTCTCTCCGAGAACCTTCTCTGTG 964
 QY 1005 AGCCTATATACCAATATGACCGGCGACATCTGGCTTTGAGGCACTTCTTCCAGCTG 1064
 Db 965 ACACATATATACCAATATGACCGGCGACATCTGGCTTTGAGGCACTTCTTCCAGCTG 1024
 QY 1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTGTAGTGAACCCCAAGGACATTTAGCAGCC 1124
 Db 1025 CTTAGGATGAGCAGCTGTGGAGGCCCTTACGTAAAGCCCAAGGACATTTCAACAGCC 1084
 QY 1125 TACTATCCAGGCCACTACCGGCCCAACATCAACTGACATGAAATCAAGGTGCCAAC 1184
 Db 1085 TACTATCCAGGCCACTACCGGCCCAACATCAACTGACATGAAATCAAGGTGCCAAC 1144
 QY 1185 AACCGGAAGTGAAGGTGGCTTCAAACTCTTCTATCTGTGTGAGCCCAACGTACAGT 1244
 Db 1145 AACCGGAAGTGAAGGTGGCTTCAAACTCTTCTATCTGTGTGAGCCCAACGTACAGT 1204
 QY 1245 GGTCTCTGACCAAGGACATATGTGAGATCAACGGGAGAAAGTACTCGGTGAGGCTC 1304
 Db 1205 GGTCTCTGACCAAGGACATATGTGAGATCAACGGGAGAAAGTACTCGGTGAGGCTC 1264
 QY 1305 CAGTTTGTGTGAGAGCAAGCAAGCAGCAAGATTAAGTCACTTCAATCTGATCACTG 1364
 Db 1265 CAGTTTGTGTGAGAGCAAGCAAGCAGCAAGATTAAGTCACTTCAATCTGATCACTG 1324
 QY 1365 TACAGGAGACCGGGTCTTCTAGCTAGTACCTCTCTCTAGACTTCCACAGCCGTTGCCA 1424
 Db 1325 TACAGGAGACCGGGTCTTCTAGCTAGTACCTCTCTCTAGACTTCCACAGCCGTTGCCA 1384
 QY 1425 GGGATGTTCAATGTGACAGCTGAGCGTGCATCCGAAAGAACTGGGCTGGAGCGCTGG 1484
 Db 1385 GGGATGTTCAATGTGACAGCTGAGCGTGCATCCGAAAGAACTGGGCTGGAGCGCTGG 1444
 QY 1485 GCAAGCTCCCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1544
 Db 1445 GCGGATGCAAGCAAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504
 QY 1545 ACGTCAAAAACCAATTTCTGCAAGCCCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1604
 Db 1505 ACGTCAAAAACCAATTTCTGCAAGCCCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1564
 QY 1605 GGGAGCGAAGTGAAGAGAGGCTGCACTGTCTGCTGGGAGTTTCAAGTGTCAAT 1664
 Db 1565 GGAAGCAACAGAGAGAGAGGCTGCACTGTCTGCTGGGAGTTTCAAGTGTCAAT 1624
 QY 1665 GGGAGCGAAGTGAAGAGAGGCTGCACTGTCTGCTGGGAGTTTCAAGTGTCAAT 1724
 Db 1625 GGAAGCAACAGAGAGAGAGGCTGCACTGTCTGCTGGGAGTTTCAAGTGTCAAT 1684
 QY 1725 GACGAGGCTTCAATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
 Db 1685 GACGAGGCTTCAATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744
 QY 1785 CAAATGAGCTCTGTCTGAGCAAGGCAACCTTATGATGATGATGATGATGATGATGATGAT 1844
 Db 1745 CTAATGAGCTCTGTCTGAGCAAGGCAACCTTATGATGATGATGATGATGATGATGATGAT 1804
 QY 1845 GATGCTTCCATGAGAAATCTGATCTGTGGCTGTGATCTTTTACCAACAGGCTGC 1904
 Db 1805 GACGCTTCAATGAGAAATCTGATCTGTGGCTGTGATCTTTTACCAACAGGCTGC 1864

```

QY 1905 GTGTTGTGTCACGATGCGGACGAGGCGAGTGCCCTTGGCAGGTGAGCTTCACGCC 1964
    |||
    |||
    |||
Db 1865 GTTGTGGGGGACGATGCGATGAGGGCGAGTGCCCTTGCGAGGTAGCCCTGCATGCT 1924
    |||
    |||
    |||
QY 1965 CTGGGCGAGGGGCACTGTGTGGGGGCTGGCTCATCTCTCTGACTGTGGCTCTGTGA 2024
    |||
    |||
    |||
Db 1925 CTGGGCGAGGGGCACTGTGGGGGCTGGCTCATCTCTCTGACTGTGGCTCTGTGC 1984
    |||
    |||
    |||
QY 2025 GCTCATTTGCTTTCAGAGTACCAAAAATTTCAAGTACTCAGACTACGATGTGGAGGCC 2084
    |||
    |||
    |||
Db 1985 GCAACACTGCTACATCATGATGACAGAGATTCAAGTACTCAGACCACGACAGTGGAGGCC 2044
    |||
    |||
    |||
QY 2085 TTTCCTGGGTCTGTGTGACACGAGCAAGCGAGTCTTGGGGTGGAGAGCTGGAAGCTC 2144
    |||
    |||
    |||
Db 2045 TTTCCTGGGTCTGTGTGACACGAGCGAGCGAGCGCCCTGGGGTGGAGAGCGAGGCTC 2104
    |||
    |||
    |||
QY 2145 AAACGATATCATACCCACCCCTTCTTCAATGATTTTCACTTTCAGACTATGACATCGCCTTG 2204
    |||
    |||
    |||
Db 2105 AAGCGCATCTCTCCCACTCCCTTCTTCAATGACTTCACTTTCAGACTATGACATCGGCTTG 2164
    |||
    |||
    |||
QY 2205 CTGGAGCTGGAGAAATCGATGAGTACAGCAGCGTGGTGGCCCATCTGCTGCTGAT 2264
    |||
    |||
    |||
Db 2165 CTGGAGCTGGAGAAACCGGACAGTACAGCTTCATGATGGTGGGCCCATCTGCTGCGGAC 2224
    |||
    |||
    |||
QY 2265 GCTAACCATATCTTCTCTGTGTGCAAGGCGCATCTGGGTGACAGGCTGGGGGCGACACAAA 2324
    |||
    |||
    |||
Db 2225 GCTCCCATATCTTCTCTGTGTGCAAGGCGCATCTGGGTGACAGGCTGGGGGCGACACCCAG 2284
    |||
    |||
    |||
QY 2325 GAGGAGAGTACCGGAGCGTGTCTTGCAGAAAGGTGAGATCCGTGTCTATCAACACGAGCC 2384
    |||
    |||
    |||
Db 2285 TAGGAGAGGACCTGGCGGCTGATCTCTCAAAAAGGTGAGATCCGCGCATCAACACGAGCC 2344
    |||
    |||
    |||
QY 2385 ACCTGTGAGACCTCATGCGCGACAGATACACCCACGAATGATGTGTGGTTTCTC 2444
    |||
    |||
    |||
Db 2345 ACCGTGAGAACCTCTGCGCGACAGATACACCGCGCATGATGTGTGGTTTCTC 2404
    |||
    |||
    |||
QY 2445 AGTGGGGGTGTGTGACTCTCTGCGAGGAGTCTGTGTGGCCCTTGTCAAGGGCGAGAAA 2504
    |||
    |||
    |||
Db 2405 AGCGGCGGCTGTGACTCTCTGCGAGGAGTCTGTGTGGCCCTTGTCAAGGGCGAGGCG 2464
    |||
    |||
    |||
QY 2505 GATGGGCGAATGTTCCAGAGCTGTGTGTGAGTGTGGGTGAGAGCTCGCTCAGAGGAGAC 2564
    |||
    |||
    |||
Db 2465 GATGGGCGGATCTTCCAGGCGCGGTGTGTGAGTGTGGGAGAGCGCTCGCTCAGAGGAGAC 2524
    |||
    |||
    |||
QY 2565 AAGCCAGAGCTGTACACAGAGCTCTGTGATTTGGGACTGTGATCAAAAGAGCACTGGG 2624
    |||
    |||
    |||
Db 2525 AAGCCAGAGCTGTACACAGAGCTCTGTGATTTGGGACTGTGATCAAAAGAGCACTGGG 2584
    |||
    |||
    |||
QY 2625 GTATAGCAGCATGACAGACGAGCCGACACAAACACCAAGAGGATGCGCGACATGCA 2684
    |||
    |||
    |||
Db 2585 GTATAGGAGGCGCGG---GCCACCCCAATGTGTACACTGCGGGGCGACCCATCTGTCAAC 2641
    |||
    |||
    |||
QY 2685 CCTGGATACAGAGAGAGAACACTGACGACATTTATGTGTGGCCCTGCCCCCGCAACACA 2744
    |||
    |||
    |||
Db 2642 CCAGTGTGCAAG-CCTGAGAGCTGGAAGCTGGAACCGCTGACATGCAACGAGCCGCC-CCAGA 2699
    |||
    |||
    |||
QY 2745 ACCCAGACTGTGAATCTGATCTTTAGACTAGAGT 2780
    |||
    |||
    |||
Db 2700 ACATACACTGTGAATCTCAATCTCCAGGGGCTCAAAAT 2735
    |||
    |||
    |||

```

RESULT 13
AAL60793
ID AAL60793 standard; DNA; 3147 BP.

XX AAL60793;

XX 03-SEP-2003 (first entry)

XX Human membrane-type serine protease MTSPI protease domain DNA.

XX Serine protease 17; CVPSP17; tumour; cancer; antisense therapy; prostate;

```

KM breast; cervix; lung; ovary; colon; gene therapy; human; enzyme; MTSPI;
KW membrane-type serine protease; matrixase; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1865..2590
XX FT /tag=a
XX FT /product="Human matrixase protease domain"
XX MO2003044179-A2.
XX 30-MAY-2003.
XX 20-NOV-2002; 2002WC-US037626.
XX 20-NOV-2001; 2001US-0332015P.
XX (CORV-) CORVAS INT INC.
XX Madison EL, Ong EO;
XX WPI; 2003-449816/42.
XX P-PSDB; AAO30147.
XX
PT New substantially purified serine protease 17 polypeptide and encoding
PT nucleic acid, useful for diagnosing and treating tumor conditions and/or
PT cancer, particularly of the breast, cervix, prostate, lung, ovary or
PT colon.
XX
XX Disclosure; Page 177-179; 189pp; English.
XX
XX The invention relates to serine protease 17 polypeptide designated CVPSP17
XX and its corresponding nucleic acid sequence. The invention also relates
XX to a method using CVPSP17 protein to identify compounds that modulate its
XX protease activity. The method is useful for preventing, diagnosing and
XX treating disorders related to the serine protease 17 activity, such as
XX tumour conditions and/or cancer, particularly of the breast, prostate,
XX cervix, lung, ovary or colon. CVPSP17 DNA is used in gene therapy and in
XX antisense therapy. The present sequence is human membrane-type serine
XX protease MTSPI (also called matrixase) protease domain DNA. This
XX sequence is used to illustrate the method of the invention
XX
SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

```

Query Match 60.6%; Score 1883.2; DB 7; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

```

QY 45 GATCGAACCGCGCAAAACCATGGGTAGCAATCGGGGCGGCAAGCGCGAGGGGGCTTCAG 104
    |||
    |||
    |||
Db 5 GAGCGGCGCTCGGGGTACCAATGGGAGCGATCGGGCGCGCAAGGGCGGAGGGGCGCGAAG 64
    |||
    |||
    |||
QY 105 GACTTGGCGCGGGGACCTCAAGTACAACTCCCGGCTTAGAGAAACATGATGGCTTGAAGAG 164
    |||
    |||
    |||
Db 65 GACTTGGCGCGGGGACCTCAAGTACAACTCCCGGCAAGAGAAAGTATGGCTTGAAGAGAA 124
    |||
    |||
    |||
QY 165 GGTGTGAGTTCCTGCGCTGCGAACAATGCCAAGAAAGTGGAGAACGAGGCCGAGCGCG 224
    |||
    |||
    |||
Db 125 GCGGTGAGTTCCTGCGCTGCGAACAATGCCAAGAAAGTGGAGAACGAGGCCGAGCGCG 184
    |||
    |||
    |||
QY 225 TGGGTGTGTGTGTGGGAGTGTCTGAGCTTCTCTTGTCTCTCTCTCTCTCTCTCTCTCT 284
    |||
    |||
    |||
Db 185 TGGGTGTGTGTGTGGGAGTGTCTGAGCTTCTCTTGTCTCTCTCTCTCTCTCTCTCTCT 244
    |||
    |||
    |||
QY 285 CTGGTGTGGCACTTCCATTATCCGAATGTGGGGTCAAAAAGTTCATGTGGCCCATCTG 344
    |||
    |||
    |||
Db 245 CTGGTGTGGCACTTTCATACCGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 304
    |||
    |||
    |||
QY 345 AGGATCACAAAATGAGATCTTCTGTGATCGATGAGAACTCCACCTCCACAGAGTTTATC 404
    |||
    |||
    |||
Db 305 AGGATCACAAAATGAGATTTTGTGATCGCTACGAGAACTCCACCTCCAGAGTTTATC 364
    |||
    |||
    |||

```

QY 405 AGCTGGCCAGCCAGTGAAGAGGCGGTGAAGCTGCTGTACATGAATGCCCTGTCTG 464
DB 365 AGCTGGCCAGCCAGTGAAGAGGCGGTGAAGCTGCTGTACATGAATGCCCTGTCTG 424
QY 465 GGTCCCTACCAAGAAAGTGGCTGTAACTGCTTCAAGTAGGGCAGTGTCAATGACCTAC 524
DB 425 GGGCCCTACCAAGAAAGTGGCTGTAACTGCTTCAAGTAGGGCAGTGTCAATGACCTAC 484
QY 525 TACTGGTCAAGTTCAGCATTCCTCCCAACCTGCGAGAAAGGTTGATCGCCCATGGCT 584
DB 485 TACTGGTCAAGTTCAGCATTCCTCCCAACCTGCGAGAAAGGTTGATCGCCCATGGCT 544
QY 585 GTGAGAGAGTGTGAATCATTCGACACCCCGAGACGGGCACTGAAATCTTCTGTACACA 644
DB 545 GAGAGAGAGTGTGAATCATTCGACACCCCGAGACGGGCACTGAAATCTTCTGTACACA 604
QY 645 TCTGTGTGGCTTCTCCCAATTGACCCCAAGATGTGAGAGAGTCAAGGACACAGCTGC 704
DB 605 TCAAGTGTGGCTTCTCCCAATTGACCCCAAGATGTGAGAGAGTCAAGGACACAGCTGC 664
QY 705 AGTTTGGCTGTGATGCCATGGTGAAGAGTGAACAGCTTCACTAACCCCTGCTTCCCT 764
DB 665 AGTTTGGCTGTGATGCCATGGTGAAGAGTGAACAGCTTCACTAACCCCTGCTTCCCT 724
QY 765 AACAGTCCCTACCCCGGCGATGCGGCTGCGAGTGGTCTGTGGGGGACGCCACTCT 824
DB 725 GACAGCCCTTACCCCGGCTGATGCGGCTGCGAGTGGTCTGTGGGGGACGCCACTCT 784
QY 825 GTGCTGAGCTTCACTTCCGAAAGCTTGTGATGCTCTCTGTGATGAGCATGGCACTGAC 884
DB 785 GTGCTGAGCTTCACTTCCGAGCTTGTGACCTTGTGCTCTGTGAGAGAGCGGCAAGAC 844
QY 885 CTGGTCACTGCTGTATATAGCTGAGACCCCATGGAACCCCATGCTGTGGCTGTGT 944
DB 845 CTGGTCACTGCTGTATATAGCTGAGACCCCATGGAACCCCATGCTGTGGCTGTGT 904
QY 945 GGCACCTTCTCAACCCCTCAACAACCTGATCTTCTCTCTCCAGAAAGTCTTCTTGTTC 1004
DB 905 GGCACCTTCTCAACCCCTCAACAACCTGATCTTCTCTCTCCAGAAAGTCTTCTTGTTC 964
QY 1005 AGCTGATTAACCAATATGACCGGCGACATCTGCTTGAAGGCACTTCTTCCAGCTG 1064
DB 965 AACTGATTAACCAATATGACCGGCGACATCTGCTTGAAGGCACTTCTTCCAGCTG 1024
QY 1065 CCCAAGATGAGAGCTGTGGCGGCTTGTGATGACACCCCAAGGACATTTAGCAGCCCT 1124
DB 1025 CCTAGGATGAGAGCTGTGGCGGCTTGTGATGACACCCCAAGGACATTTAGCAGCCCT 1084
QY 1125 TACTATCCAGGCGACATACCCCGCCCAATCACTGCACTGGAATATCAAGGTGCCAAC 1184
DB 1085 TACTATCCAGGCGACATACCCCGCCCAATCACTGCACTGGAATATCAAGGTGCCAAC 1144
QY 1185 AACCGGAACGTGAAGTGGCTTCAAACTCTTCTTCTGTGGAGCCCAAGTACCAAGT 1244
DB 1145 AACCGGAACGTGAAGTGGCTTCAAACTCTTCTTCTGTGGAGCCCAAGTACCAAGT 1204
QY 1245 GGCCTCTCCACCAAGAGCTATGTGAGATCAACCGGAGAAATCTGCGGTGAGAGCTCC 1304
DB 1205 GGCCTCTCCACCAAGAGCTATGTGAGATCAACCGGAGAAATCTGCGGTGAGAGCTCC 1264
QY 1305 CAGTTTGTGAGAGAGCAAGCAAGCAAGATTAAGTCACTTCAATCTGTATACCTG 1364
DB 1285 CAGTTTGTGAGAGCAAGCAAGCAAGATTAAGTCACTTCAATCTGTATACCTG 1324
QY 1365 TACAAGGACACCGGCTTCTAGCTAGTAACTCTCTTCAAGTCTCAAGAGACCGGCTCCA 1424
DB 1325 TACAAGGACACCGGCTTCTAGCTAGTAACTCTCTTCAAGTCTCAAGAGACCGGCTCCA 1384
QY 1425 GGGATTTTCAATGTGAGAGAGTGAAGTGTGATCCGAAAGAACTGCGTGGACGCGCTGG 1484
DB 1385 GGGATTTTCAATGTGAGAGAGTGAAGTGTGATCCGAAAGAACTGCGTGGACGCGCTGG 1444
QY 1485 GCAGACTGCGCGGATTAATGATGATGAGCTTACTGCGCATGCAATGCCACCAAGTTCC 1544

DB 1445 GCCGACTGACCCGACCAACAGCATGAGCTCAACTGCAAGTTGGCAACCGCGGCAACCAATTC 1504
QY 1545 ACGTGCAGAAAACCAAGTTTTCGAAAGCCCTTCTTGGCTCTGTACAGTGTCAACGACTGT 1604
DB 1505 ACGTGCAGAAAACCAAGTTTTCGAAAGCCCTTCTTGGCTCTGTACAGTGTCAACGACTGT 1564
QY 1605 GGGAGCGAAGTGAAGAGAGGAGCTGAGCTGTCTGTGGAGTTTCAAGTGTCCAAAT 1664
DB 1565 GGGAGCGAAGTGAAGAGAGGAGCTGAGCTGTCTGTGGAGTTTCAAGTGTCCAAAT 1624
QY 1665 GGGAGCTGTCTCCCTCAAGAGCCCAAGATGTATGAGAGAGCAACTGTGAGATGGGCTCT 1724
DB 1625 GGGAGCTGTCTCCCTCAAGAGCCCAAGATGTATGAGAGAGCAACTGTGAGATGGGCTCT 1684
QY 1725 GAGAGAGCTTCAATGTGAACAGCTGAAATGTGTCTCTTTCACCAATATATCAACGCTGC 1784
DB 1685 GAGAGAGCTTCCCTCCCAAGGAGTGAACGTGTCACTTGTATCAAAACCACTACCGCTGC 1744
QY 1785 CAAATATGAGCTCTGTCTGACCAAGGCGCAACCTGAGTGTGATGGAGAACGCACTGTAGC 1844
DB 1745 CTCAATGAGCTCTGTCTGACCAAGGCGCAACCTGAGTGTGATGGAGAACGCACTGTAGC 1804
QY 1845 GATGCTCTGATGAGAAAACTGTGACTGTGGCTGCGATCTTTTACCAACAGCTGCG 1904
DB 1805 GATGCTCTGATGAGAAAACTGTGACTGTGGCTGCGATCTTTTACCAACAGCTGCG 1864
QY 1905 GTGGTGTGTGAGCAAGATGCGGACGAGGCGAGTGGCCCTGCGAGTGAAGCTTCAAGCC 1964
DB 1865 GTGGTGTGTGAGCAAGATGCGGACGAGTGGCCCTGCGAGTGAAGCTTCAAGCT 1924
QY 1965 CTGGGCGAGGCGCACTGTGTGTGGGCTGCTCATCTCTGACTGGCTGGCTGTGCA 2024
DB 1925 CTGGGCGAGGCGCACTGTGTGTGGGCTGCTCATCTCTGACTGGCTGGCTGTGCA 1984
QY 2025 GCTCATTTGCTTCAAGATGACAAAAATTTCAAGTACTCAGACTACAGATGTGAGAGGCG 2084
DB 1985 GCTCATTTGCTTCAAGATGACAAAAATTTCAAGTACTCAGACTACAGATGTGAGAGGCG 2044
QY 2085 TTCTGGGCTGTGCTGAGACCAAGGCAAGGCGAGTGGCTCTGGGGGTGAGAGCTGAAGCTTC 2144
DB 2045 TTCTGGGCTGTGAGACCAAGGCGAGTGGCTCTGGGGGTGAGAGCTGAAGCTTC 2104
QY 2145 AAACGATATCAATCAACCACTCTCTTCAATGATTTCACTTGAATGATGATGATGATGATG 2204
DB 2105 AAACGATATCAATCAACCACTCTCTTCAATGATTTCACTTGAATGATGATGATGATGATG 2164
QY 2205 CTGAGCTGAGAAAGTGGTGAAGTGAAGCAACCGTGTGCGCCCATCTGCTGCTGAT 2264
DB 2165 CTGAGCTGAGAAAGTGGTGAAGTGAAGCAACCGTGTGCGCCCATCTGCTGCTGAT 2224
QY 2265 GCTATCCCATGCTTCTTCCCTGCTGAGCAAGGCGATCTGGGTGACAGGCTGGGGGCAACAAA 2324
DB 2225 GCTATCCCATGCTTCTTCCCTGCTGAGCAAGGCGATCTGGGTGACAGGCTGGGGGCAACAAA 2284
QY 2325 GAGGAGGTATCCGAGGCGCTGATCTCTGCAAGAGGTTGATCCGTTGATCAACACAGAGC 2384
DB 2285 GAGGAGGTATCCGAGGCGCTGATCTCTGCAAGAGGTTGATCCGTTGATCAACACAGAGC 2344
QY 2385 AACTGTGAGACCTTCATGCGCGAGAGATCAACCCCAAGATGATGTGTGTGGTGTCTTC 2444
DB 2345 AACTGTGAGACCTTCATGCGCGAGAGATCAACCCCAAGATGATGTGTGTGGTGTCTTC 2404
QY 2445 AGTGGGGTGTGAGCTCTGCGCAAGGTTGATCTGTGTGGCCCTTGTCAAGGCGGAGAAA 2504
DB 2405 AGTGGGGTGTGAGCTCTGCGCAAGGTTGATCTGTGTGGCCCTTGTCAAGGCGGAGAAA 2464
QY 2505 GATGGGCGAATGTTTCAAGGCTGTGTGTGAGCTGGGTGAGAGGCTGGCTCAAGAGAGAC 2564
DB 2465 GATGGGCGAATGTTTCAAGGCTGTGTGTGAGCTGGGTGAGAGGCTGGCTCAAGAGAGAC 2524
QY 2565 AAGCGAGGCTGTACAAAGGCTCTGTGATTTGGGACCTGATCAAAAGACCAACTGGG 2624

Db 2525 AAGCCAGGCGGTGTACCAAGGCTCCTCTGTTTGGGACTGTGATCAAGAAGAACTGGG 2584
QY 2625 GTATAGCAGCATGAGACAGACGCGACCAACACCCAGGAGATGCCGACATGCACA 2684
Db 2585 GTATAGGCGCCGGG--GCCACCCCAATGTGTACCTGCGGGGCCACCCATGTCTCAC 2641
QY 2685 CCTGTATACAGAGAGAGAACTGAGACATTTATGCTGTGGCTCCCCCCCCCAACACA 2744
Db 2642 CCAAGTGTGACG-CCTGCAAGGCTGAGACTGACCGCTGACTGACACGAGCCCC-CCAGA 2699
QY 2745 ACCGAGCTGTGAATGACATCTCTTAGAGCTCAGAGT 2780
Db 2700 ACATACACTGTGAATCAATCTCCAGGCGCTCCAAAT 2735

RESULT 14
AAL60792 standard; DNA; 3147 BP.
AAL60792:
03-SBP-2003 (first entry)
Human membrane-type serine protease MTSPI DNA.
Serine protease 17, CVPSP17; tumour; cancer; antisense therapy; prostate;
breast; cervix; lung; ovary; colon; gene therapy; human; enzyme; MTSPI;
membrane-type serine protease; matrilysin; gene; ds.
Homo sapiens.
Location/Qualifiers
Key 23..2599
CDS /*tag= a
FT /product= "Human matrilysin"
FT
XX MO2003044179-A2.
XX 30-MAY-2003.
XX 20-NOV-2002; 2002WO-US037626.
XX 20-NOV-2001; 2001US-0332015P.
XX (CORV-) CORVAS INT INC.
XX Madison EL, Ong EO;
XX WPI; 2003-449816/42.
XX P-PSDB; AMO30146.
XX
XX New substantially purified serine protease 17 polypeptide and encoding
XX nucleic acid, useful for diagnosing and treating tumor conditions and/or
XX cancer, particularly of the breast, cervix, prostate, lung, ovary or
XX colon.
XX
XX Disclosure; Page 171-175; 189pp; English.
XX
XX The invention relates to serine protease 17 polypeptide designated CVPSP17
XX and its corresponding nucleic acid sequence. The invention also relates
XX to a method using CVPSP17 protein to identify compounds that modulate its
XX protease activity. The method is useful for preventing, diagnosing and
XX treating disorders related to the serine protease 17 activity, such as
XX tumour conditions and/or cancer, particularly of the breast, prostate,
XX cervix, lung, ovary or colon. CVPSP17 DNA is used in gene therapy and in
XX antisense therapy. The present sequence is human membrane-type serine
XX protease MTSPI (also called matrilysin) DNA. This sequence is used to
XX illustrate the method of the invention
XX
XX Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;
XX
XX Query Match 60.6%; Score 1883.2; DB 7; Length 3147;
XX Best Local Similarity 81.2%; Pred. No. 0;

Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
QY 45 GATTCGACCGCCAAAACCATGGGTAGCATCGGCGCCGAGCGCGAGGCGCTCTCAG 104
Db 5 GAGCGGCTCGGGGATGACATGGGAGACCATCGGCCCCGAGCGGAGCGGCGCCGAGG 64
QY 105 GACTTCGCGCGCGGAGCTCAAGTACAACTCCCGCTAAGAGAACATGATGCTTTGAGAG 164
Db 65 GACTTCGCGCGGAGCTCAAGTACAACTCCCGCAGAGAAATGATGCTTTGAGAGAA 124
QY 165 GGTGTGAGATTCTGCTCCGAAACATGCGAAGAGTGAAGAGGAGGAGCGCCAGCGCG 224
Db 125 GCGGTGAATTCCTGCAATCAACAGTCMAAGGTGAAGAGAGAGAGCGCGCGCG 184
QY 225 TGGGTGTGCTGTGTCAGAGTGTCTGCTCTCTGCTCTCTCATGAGTGTGCTG 284
Db 185 TGGGTGTGCTGTGTCAGAGTGTCTGCTCTCTGCTCTCTCATGAGTGTGCTG 244
QY 285 CTGCTGTGCACTTCATTAATGGAATGTGCGGCTTCAAAAAGTCTTCAATGGCCATCTG 344
Db 245 CTGCTGTGCACTTCATTAATGGAATGTGCGGCTTCAAAAAGTCTTCAATGGCCATG 304
QY 345 AGCATCAACAAATGAGATCTTTCTGATGCGGTATGAGAACTCCACTCAGAGATTATC 404
Db 305 AGCATCAACAAATGAGATTTTGTGATGCTTACGAACTTCAACTCAGAGTTTGTGA 364
QY 405 AGCCTGCGCAGCCAGGTGAAGAGAGCGCTGAGCTGTGTACATGAAGTCTCTGCTG 464
Db 365 AGCCTGCGCAGCAAGGTGAAGAGAGCGCTGAGCTGTGTACAGCGGAGTCTTCAATCTCTG 424
QY 465 GGTCTTACCAAGAAAGTGTGCTGTATCTGCTTCACTGATGAGGCACTGATCGCTAC 524
Db 425 GGCCTTACCAAGAAAGTGTGCTGTATCTGCTTCACTGATGAGGCACTGATCGCTAC 484
QY 525 TACTGTCAAGTTCAGCACTCCCGCCCACTGCGCAAGAGTGTATGCGGCTATGCT 584
Db 485 TACTGTCAAGTTCAGCACTCCCGCCCACTGCGCAAGAGTGTATGCGGCTATGCT 544
QY 585 GTGAGCGAGTGTGAATTCAGCACTCCCGCCCACTGCGCAAGAGTGTATGCGGCTATG 644
Db 545 GAGAGCGGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
QY 645 TCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 704
Db 605 TCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664
QY 705 AGTTTGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
Db 665 AGTTTGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724
QY 765 AACAGTCTTACCGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 824
Db 725 GACAGCCCTTACCGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784
QY 825 GTGCTGAGCTTACCTTCCGAGCTTTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
Db 785 GTGCTGAGCTTACCTTCCGAGCTTTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844
QY 885 CTGTGACCCGTTGATGATGCTGAGCTGAGCCCAAGAAACCCAGCTGTGTGCTGTGCTG 944
Db 845 CTGTGAGCGGTGATGATGCTGAGCTGAGCCCAAGAAACCCAGCTGTGTGCTGTGCTG 904
QY 945 GGCACCTTCTACCTCTCTTACACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004
Db 905 GGCACCTTCTACCTCTCTTACACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 964
QY 1005 AGCTGTATACCAATATCTGACCGGAGAACTCTGCTTGTGAGCCACTTCTTCCAGCTG 1064
Db 965 ACACGTATACCAACATGAGGCGGAGCTTCCGCTTGTGAGCCACTTCTTCCAGCTG 1024
QY 1065 CCAAGATGAGAGCTGTGCGGCTTTTGTAGTGAACCCAGAGGACATTTAGACGCC 1124
Db 1025 CCAAGATGAGAGCTGTGCGGCTTTTGTAGTGAACCCAGAGGACATTTAGACGCC 1084

QY 1125 TACTATCCAGGCGCTACCTCCGCAAGATCACTGCACTATGGAAATATCAAGTGTCCCAAC 1184
 DB 1085 TACTATCCAGGCGCTACCTCCGCAAGATCACTGCACTATGGAAATATCAAGTGTCCCAAC 1144
 QY 1185 AACCGAAGCTGGAAGTGTGCGCTTCAAACTCTTATCTGAGTGAAGCCCAAGTACAGTGC 1244
 DB 1145 AACCGAAGCTGGAAGTGTGCGCTTCAAACTCTTATCTGAGTGAAGCCCAAGTACAGTGC 1204
 QY 1245 GGTCTCTGCAACCAAGACTATGTGAGATCAACCGGGGAGAAGTACTGCGGTGAAGAGTTC 1304
 DB 1205 GGGACCTGCCCCCAAGGACTACGTGAGATCAATGGGGAAGAAATATCTGCGAGAGAGTTC 1264
 QY 1305 CAGTTGTGTGAGGAGCAACAGAGCAAGATTAAGTCACTTCCATTTGATCACTGC 1364
 DB 1265 CAGTTGTGTGAGGAGCAACAGAGCAAGATTAAGTCACTTCCATTTGATCACTGC 1324
 QY 1365 TACACGCAACCGGGTCTCTAGCTGAGTACTCTCCCTAGCACTCCCAAGCCGCTGCCA 1424
 DB 1325 TACACGCAACCGGGTCTCTAGCTGAGTACTCTCCCTAGCACTCCCAAGCCGCTGCCA 1384
 QY 1425 GGGATGTTCACTGTGCAAGCTGACCGGTGCATCCGAAAGGAACTGCGCTGCGACGCGTGC 1484
 DB 1385 GGGATGTTCACTGTGCAAGCTGACCGGTGCATCCGAAAGGAACTGCGCTGCGACGCGTGC 1444
 QY 1485 GCAGACTGCGCGGATTAATGATGATGAGCGCTTACTGCGCATGCAATGCCACCCACGCTTC 1544
 DB 1445 GCGGACTGCAACCGACCAACGAGTGAAGTCACTGACGTTGCAACCCCGGCAACGAGTTC 1504
 QY 1545 ACGTGGAAAAACCACTTCTGCAAGCCCTCTTCTGGGCTCTGAGCAAGTGTCAACGACTGC 1604
 DB 1505 ACGTGGAAAAACCACTTCTGCAAGCCCTCTTCTGGGCTCTGAGCAAGTGTCAACGACTGC 1564
 QY 1605 GGGGACGGAAGTGAACGAGAGGCGCTGACGCTGTCTGCGGAGTTCAAGTGTCCAAAT 1664
 DB 1565 GGGGACGGAAGTGAACGAGAGGCGCTGACGCTGTCTGCGGAGTTCAAGTGTCCAAAT 1624
 QY 1665 GGGAAAGTCTGCGCTTCAAGGCGCAAGTGAATGGAAGAGCAACTGTGGAAGTGGGTCT 1724
 DB 1625 GGGAAAGTCTGCGCTTCAAGGCGCAAGTGAATGGAAGAGCAACTGTGGAAGTGGGTCT 1684
 QY 1725 GACGAGGCTTCACTGATGACGCGTGAATGCTGCTCTTGGACCAAAATATACCTACCGCTGC 1784
 DB 1685 GACGAGGCTTCACTGATGACGCGTGAATGCTGCTCTTGGACCAAAATATACCTACCGCTGC 1744
 QY 1785 CAAAATGCGCTCTGTCTGAGCAAGGCAACCTGAGTGTGATGGAGAGCGACTGTAGC 1844
 DB 1745 CTAATGCGCTCTGTCTGAGCAAGGCAACCTGAGTGTGATGGAGAGCGACTGTAGC 1804
 QY 1845 GATGCTCCGATGAGAAAACTGTGACTGTGGGCTGCGATCCTTTTCCAAACAGGCTGC 1904
 DB 1805 GATGCTCCGATGAGAAAACTGTGACTGTGGGCTGCGATCCTTTTCCAAACAGGCTGC 1864
 QY 1905 GTGTTGTGTGACCAATGCGGACGAGAGGCGAGTGGGCTCTGACAGTGGAGCTCCAGCGC 1964
 DB 1865 GTGTTGTGTGACCAATGCGGACGAGAGGCGAGTGGGCTCTGACAGTGGAGCTCCAGCGC 1924
 QY 1965 CTGGGCTCAGGCGCACTTGTGTGGGCTGCTCATCTCTCTGACTGGTGGTCTCTGCA 2024
 DB 1925 CTGGGCTCAGGCGCACTTGTGTGGGCTGCTCATCTCTCTGACTGGTGGTCTCTGCA 1984
 QY 2025 GCTCATTTGCTTCAAGATGAGCAAAATTTCAAGTACTCAAGTACACGATGTGAGCGGC 2084
 DB 1985 GCAACCTCTCATGATGAGCAAGGATTCAGTACTAGAACCCACGACGAGTGGAGCGGC 2044
 QY 2085 TTCTGGGTCTGCTGAGCAAGGCAAGGAGTGGCTCTGAGGAGTGGAGAGTGGAGGCTC 2144
 DB 2045 TTCTGGGTCTGCTGAGCAAGGCAAGGAGTGGCTCTGAGGAGTGGAGAGTGGAGGCTC 2104
 QY 2145 AAAGCATATCATCCACCTCTTCTTCAATGATTTCACTTGAAGTATGAGACTGAGCTG 2204
 DB 2105 AAAGCATATCATCCACCTCTTCTTCAATGATTTCACTTGAAGTATGAGACTGAGCTG 2164

QY 2205 CTGAGAGCTGAGAAAGTGTGAGAGTACAGACCGCTGCGGCCCATCTGCTCTGAT 2264
 DB 2165 CTGAGAGCTGAGAAAGTGTGAGAGTACAGACCGCTGCGGCCCATCTGCTCTGAT 2224
 QY 2265 GCTACCATATCTTCTCTGCTGAGAGGCGCATCTGGGTCAACAGCTGGGGGCAACAAAA 2324
 DB 2225 GCTACCATATCTTCTCTGCTGAGAGGCGCATCTGGGTCAACAGCTGGGGGCAACAAAA 2284
 QY 2325 GAGGAGGCTACCGGAGCGCTGATCTCTGACAGAGGAGTATCCGTGTCAACCAACACC 2384
 DB 2285 TATGAGAGCACTGGCGGCTGATCTGCAAAAGGAGTATCCGTGTCAACCAACACC 2344
 QY 2385 ACCGTGAGGACCTCATGCTGCGGACAGATCAACCCAGATGATGTGTGGTTCCTC 2444
 DB 2345 ACCGTGAGGACCTCATGCTGCGGACAGATCAACCCAGATGATGTGTGGTTCCTC 2404
 QY 2445 AGTGGGGGTGTGAGCTCTGCGAGGAGTACTGTGTGCGCCCTTGTCAACCGGAGAAAA 2504
 DB 2405 AGTGGGGGTGTGAGCTCTGCGAGGAGTACTGTGTGCGCCCTTGTCAACCGGAGAAAA 2464
 QY 2505 GATGGGCGAATGTTCCAGGCTGTGTGTGAGTGTGAGGCTGTGAGAGCTGCTCAGAGAAC 2564
 DB 2465 GATGGGCGAATGTTCCAGGCTGTGTGTGAGTGTGAGGCTGTGAGAGCTGCTCAGAGAAC 2524
 QY 2565 AAGCCAGGCGGTGACCAAGAGCTCCCTGTAGTTCGGGACTGATCAAAAGACACACTGG 2624
 DB 2525 AAGCCAGGCGGTGACCAAGAGCTCCCTGTAGTTCGGGACTGATCAAAAGACACACTGG 2584
 QY 2625 GTTATGACAGATGAGACAGACAGCCGACCAACCAACCCACAGGAGTCCGACATGCACA 2684
 DB 2585 GTTATGACAGATGAGACAGACAGCCGACCAACCAACCCACAGGAGTCCGACATGCACA 2644
 QY 2685 CCTGATACAGAGAGAGAACATGACGACATTTATGCTGTGGCTGCCCCCCCCCAACACA 2744
 DB 2642 CCAAGTGTGACG-CCTGACAGCTGAGACTGACCGCTGATCACTGACCAAGCGCC-CCAGA 2699
 QY 2745 ACCCAAGCTGTGAATGATCTCTTGAAGTCACTGAGT 2780
 DB 2700 ACATACACTGTGAATGATCTCTTGAAGTCACTGAGT 2735

RESULT 15
 AAD47225 strand; DNA; 3147 BP.
 ID AAD47225;
 AC AAD47225;
 DT 24-FEB-2003 (first entry)
 XX
 DE Human membrane-type serine protease 1 (MTSP1) DNA.
 XX
 KW Human; type II membrane-type serine protease 9; tumour; transgenic;
 KW type II transmembrane serine protease; enzyme; gene therapy; MTSP9;
 KW neoplastic disease; transgenic animal; membrane-type serine protease 1;
 KW TTSP; MTSP1; matrixinase; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 23..2590
 FT /tag=a
 FT /product="Human MTSP1 protein"
 PN MO20027267-A2.
 XX
 PD 03-OCT-2002.
 XX
 PE 27-MAR-2002; 2002WC-US009611.
 XX
 PR 27-MAR-2001; 2001US-0279228P.
 XX
 PR 15-MAY-2001; 2001US-0291501P.
 XX
 PA (CORV-) CORVAS INT INC.

XX Madison EL, Ong EO;
XX WPI; 2003-018940/01.
XX P-PSDB; AAE29820.
XX
XX New substantially purified single or two-chain type II membrane-type
XX serine protease 9 (MTSP9) polypeptide; useful for monitoring tumor
XX progression, inhibiting tumor initiation, or treating a malignant or pre-
XX malignant condition.
XX
XX Disclosure; Page 179-183; 199pp; English.
XX
XX The invention relates to type II membrane-type serine protease 9 (MTSP9)
XX polypeptides and polynucleotides. MTSP belongs to type II transmembrane
XX serine protease (TSP) family. Sequences of the invention and their
XX antibodies are useful for diagnosing, treating or preventing neoplastic
XX disease in mammals. They are useful for monitoring tumor progression,
XX inhibiting tumor initiation, growth or progression or treating malignant
XX or pre-malignant conditions. Transgenic animals of the invention are
XX useful in animal models of tumor initiation, growth and/or progression
XX models. The invention is also useful in gene therapy. The present
XX sequence is human membrane-type serine protease I (MTSP1) DNA. MTSP1 also
XX referred as matrilipase is a member of the TSP family
XX
SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other:
Query Match 60.6%; Score 1893.2; DB 7; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
QY 45 GATCGACCCGCCAAACCATGGGTAGCAATCGGGGGCGCAAGGCGGAGGGGCTCTCAG 104
DB 5 GACCGGCGCTGGGGGTACCATGAGGAGCATGGGCCCGCAAGGGCGAGGGGCCGAG 64
QY 105 GACTTCGCGCGGGGACTCAAGTACAACTCCCGCTAGAGAAATGAGTGGCTTTGAGAG 164
DB 65 GACTTCGCGCGGGGACTCAAGTACAACTCCCGCGACAGAAATGAGTGGCTTTGAGAG 124
QY 165 GGTGTGAGATTCTCGCTCGGACCAATGCGCAAGAAATGAGAAAGCGAGGCCAGCGCG 224
DB 125 GGCCTGAGATTCTCGCGAGTCAACAGTCMAAGAGGTGMAAAGCATGGGCCCGGGCGC 184
QY 225 TGGGTGTGCTGT 284
DB 185 TGGGT 244
QY 285 CTGGT 344
DB 245 CTGGT 304
QY 345 AGGATCAAAATGAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 404
DB 305 AGGATCAAAATGAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 364
QY 405 AGGCTGGCGACCGAGTGAAGAGGCGCTGAAGCTGTACAAATGAGAGTCCCTGTCTGT 464
DB 365 AGCTGGCGACCGAGTGAAGAGGCGCTGAAGCTGTACAAATGAGAGTCCCTGTCTGT 424
QY 465 GGTCCCTACCAAGAAAGTGGCTGTAACTGTCTTCACTGAGAGGCGATGTCTGTCTAC 524
DB 425 GGGCCCTACCAAGAAAGTGGCTGTAACTGTCTTCACTGAGAGGCGATGTCTGTCTAC 484
QY 525 TACTGTGTGAGATTGAGATCCCGCAGACCTGTGTGAGAGGCGAGCGCTCTATGTGCGC 584
DB 485 TACTGTGTGAGATTGAGATCCCGCAGACCTGTGTGAGAGGCGAGCGCTCTATGTGCGC 544
QY 585 GTGAGCGAGTTGTAACTTGTGCAACCGGAGCAGCGGAGCTGAATCTCTGTGTCTAACA 644
DB 545 GAGGAGCGCGAGT 604
QY 645 TCTGT 704

DB 605 TCAGT 664
QY 705 AGTTTGT 764
DB 665 AGCTTTGT 724
QY 765 AACAGT 824
DB 725 GACAGCGCTTACCGCTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 784
QY 825 GT 884
DB 785 GT 844
QY 885 CTGGT 944
DB 845 CTGGT 904
QY 945 GGCACCTTGT 1004
DB 905 GGCACCTTGT 964
QY 1005 AGCTGT 1064
DB 965 ACAGT 1024
QY 1065 CCCAAGATGAGCAGCTGT 1124
DB 1025 CTGT 1084
QY 1125 TACTATTCAGGCGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1184
DB 1085 TACTATTCAGGCGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1144
QY 1185 AACCGAAGT 1244
DB 1145 AACCGAAGT 1204
QY 1245 GGTCTGT 1304
DB 1205 GGTCTGT 1264
QY 1305 CAGTTTGT 1364
DB 1265 CAGTTTGT 1324
QY 1365 TACACGAGCAGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1424
DB 1325 TACACGAGCAGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1384
QY 1425 GGGATGT 1484
DB 1385 GGGATGT 1444
QY 1485 GCAAGT 1544
DB 1445 GCGAATGT 1504
QY 1545 ACCTGT 1604
DB 1505 ACCTGT 1564
QY 1605 GGGAGCGAAGT 1664
DB 1565 GGGAGCGAAGT 1624
QY 1665 GGGAGT 1724
DB 1625 GGGAGT 1684
QY 1725 GACGAGGCTTGT 1784
DB 1685 GACGAGGCTTGT 1744

BLANK
PAGE
